UNIVERSITY OF LIVERPOOL

Characterisation of transcription factors with potential roles in the circadian optimisation of Crassulacean acid metabolism in *Kalanchoë fedtschenkoi.*

Thesis submitted in accordance with the requirements to the University of Liverpool for the degree of Doctor of Philosophy.

Jade Louise Waller

September 2015

This PhD thesis is dedicated to my three wonderful grandparents, who have given me so many beautiful memories to treasure.

They are loved and missed everyday.

Vera Norma Bowen

Margaret Waller

Arthur William Waller, J.P, M.B.E.

Abstract.

Characterisation of transcription factors with potential roles in the circadian optimisation of Crassulacean acid metabolism in *Kalanchoë fedtschenkoi.*

Crassulacean acid metabolism (CAM) is a metabolic adaptation of photosynthesis that is optimised via strict temporal regulation of its biochemistry by the circadian oscillator. CAM plants achieve high water use efficiency and thus thrive in seasonally dry regions unsuitable for C_3 food crops such as rice or wheat. Climate change and the associated challenges of global food and energy security, mean that CAM research is currently of urgent and pressing need, as CAM may reveal methods for the generation of more water use efficient crops.

The efficiency of CAM is optimised by the circadian clock, through the regulation of *PHOSPHOENOLPYRUVATE CARBOXYLASE KINASE* (*PPCK*) expression and nocturnal CO_2 fixation, but the signalling pathway between the central clock and CAM has yet to be elucidated.

Whole genome sequencing and detailed RNA-seq datasets for C_3 and CAM leaves of the model CAM species *Kalanchoë fedtschenkoi* have enabled the discovery of novel genes that could function to link CAM to the circadian clock. Three CAM-induced and clock-controlled transcription factor (TF) genes were identified from the RNA-seq datasets: *MYB-LIKE 439* (*KfMYB439*), *CAM-INDUCED BZIP1* (*KfCIB1*) and *CYCLING DOF FACTOR2* (*KfCDF2*).

Both over-expressor and RNAi knockdown transgenic lines of *K. fedtschenkoi* were generated for each TF, facilitating the further elucidation of their biological functions. Over 200 transgenic lines were screened for altered expression levels, and changes to the dawn and dusk levels of key CAM metabolites: malate and starch. Four transgenic lines for each TF, two over-expressor and two RNAi lines, were used for detailed phenotypic analysis of CAM-associated traits.

Transgenic perturbation of any one of the three TFs caused small but widespread changes to the transcript levels of core clock- and CAM-associated genes.

KfMYB439 is a REVEILLE family TF related to the clock gene *CIRCADIAN CLOCK ASSOCIATED1 (CCA1*). Data revealed that *KfMYB439* functioned close to the core circadian oscillator. Mis-regulation of *KfMYB439* led to the perturbation of efficient dark CO_2 fixation, reduced levels of starch and malate, and reduced productivity during drought.

KfCIB1, was found to feed back to influence the core circadian clock as well as regulating CAM. In constant light conditions, *KfCIB1* mis-expression led to perturbed timing of *KfCCA1* and *TIMING OF CAB1* (*KfTOC1*) . *KfCIB1* mis-regulation also impacted on stomatal control. At dusk and dawn, large and rapid changes in stomatal conductance resulted in spikes of CO_2. Mis-expression also resulted in small improvements in productivity in water-limited environments.

KfCDF2 was found to function not only in the clock control of CAM, but also in the photoperiodic control of flowering time. In terms of CAM and the clock, *KfCDF2* mis-expression caused changes to *CCA1, TOC1* and *PPCK* expression, and arrhythmic CO_2 fixation in constant conditions. It also impacted on water retention during drought, with both over-expressor and RNAi lines displaying higher succulence than the wild type lines after 90 days of drought. *KfCDF2* over-expression in both *K. fedtschenkoi* and *K. laxiflora* caused constitutive flowering in long days, whereas wild type plants never flowered. Q-RT-PCR analysis of flowering pathway genes revealed that KfCDF2 over-expression impacted on transcript levels for *CONSTANS* (*CO*) and *FLOWERING LOCUS T* (*FT*); key proteins in the day-length dependent induction of flowering.

Results suggested that all three TFs likely function in the circadian optimisation of CAM, although whether or not the often small effects were direct or indirect will require further work. Future Chromatin Immunoprecipitation and sequencing experiments will reveal the target genes regulated by these TFs, and the identification of other novel CAM-induced genes from the RNA-seq data, will allow the pathway between the circadian clock and CAM to be elucidated in much greater detail.

Table of contents

List of Figures.

Chapter 1 -

Chapter 3 –

Chapter 4-

List of Tables

Chapter 5 -

Chapter 6 -

Glossary of abbreviations:

AG	AGAMOUS
AGL24	AGAOUS-LIKE24
ANOVA	Analysis of Variance
AP1	APETALLA1
AP2	APETALLA2
AS	Alternative Splicing
CAM	Crassulacean acid metabolism
CCA1	CIRCADIAN CLOCK ASSOCIATED1
CDF2	CYCLING DOF FACTOR2
CIB1	CAM-INDUCED bZIP1
CO	CONSTANS
CO_2	Carbon Dioxide
dH_2O	Distilled water
DOF	DNA-BINDING WITH ONE FINGER
DW	Dry Weight
ED	Evening complex
EE	Evening element
ELF3/4	EARLY FLOWERING3/4
ERD	EARLY RESPONSIVE TO DEHYDRATION
FL_OX	Full-length Overexpresser
FM	Floral Meristem
FT	FLOWERING LOCUS T
FUL	FRUITFULL
FW	Fresh Weight
GI	GIGANTEA
GPT2	GLUTAMATE-PYRUVATE TRANSAMINASE2
GWD	GLUCAN WATER DIKINASE
HA	Haemagglutinin
IM	Inflorescence Meristem
IRGA	Infra-Red Gas Analyser
K_i	Inhibition Constant
LCL domain	LHY/CCA1-LIKE domain
LD	Light/Dark
LDP	Long day flowering plant
LFY	LEAFY
LHY	LONG ELONGATED HYPOCOTYL
LL	Constant Light
ME	Morning element
MEX1	MALTOSE EXPORTER1
NAD	Nicotinamide Adenine Dinucleotide
NAD-Me α	NAD-MALIC ENZYME ALPHA
NAD-ME β	NAD-MALIC ENZYME BETA
O_2	Oxygen

PEPC	PHOSPHOENOLPYRUVATE CARBOXYLASE
PI	PISTILLA
PIF	PHYTOCHROME-INTERACTING FACTOR
PP2A	PROTEIN PHOSPHATASE TYPE 2A
PPCK	PHOSPHOENOLPYRUVATE CARBOXYLASE KINASE
PPDK	PYRUVATE ORTHOPHOSPHATE DIKINASE
PPDK-RP	PPDK REGULATORY PROTEIN
PRR	PSEUDO RESPONSE REGULATOR
RNAi	RNA interference
RuBisCO	Ribulose 1.5-bisphosphate carboxylase/oxygenase
RVE	REVEILE
SAM	Shoot Apical Meristem
SDP	Short day flowering plant
SEP	SEPALLATA
SHY2	SHORT HYPOCOTYL2
SOC1	SUPRESSOR OF OVEREXPRESSION OF CONSTANS1
TF	Transcription factor
TFL1	TERMINAL FLOWER1
TOC1	TIMING OF CAB EXPRESSION1
Tukey HSD	Tukey High Significant Difference Two Way ANOVA
umol	micro mole
WT	Wild Type
WUE	Water use efficiency
ZTL	ZEITLUPE

Acknowledgements.

Firstly I would like to thank my supervisor, Dr James Hartwell, who has given me so much help and support over these last five years, and has been key in my development as a researcher. Thank you for making your enthusiasm for CAM and plant biology so infectious!
I would also like to say a huge thank you to Louisa Dever, Susanna Boxall and Nirja Kadu who over the past few years have been so supportive, encouraging and helpful. I am so grateful for you all for being such fantastic colleagues and friends. I also feel I owe thanks to the Hartwell family for letting me adopt their Guinea Pig, and would like to thank the Head family for coming to cheer me on at the 10K tunnel run! Dave appreciates it! I must also extend my gratitude to Jean Wood, you are a wonderful person and Lab G would not be able to function without you. I would also like to thank the BBSRC for their funding and the University of Liverpool for being such a lovely place to study.

I am eternally grateful to my friends at University who have kept me motivated throughout times when nothing seemed to be working, and for celebrating with me when things finally did! From the bottom of my heart thank you.
I would especially like to thank two of my best friends, Dr Jane Pulman and Dr Chris Lofthouse, for their kind words, giggles and for the long chats. I couldn't have made it here without you both! Ben Murray, I am so proud of you getting such a fantastic job in LA (and I cannot wait to come and visit). Thanks for the sarcasm, the geography lessons, helping me 'breakout' and for turning me into a meme! Also, thank you to Ewan Young for making our office a much more interesting place to work. Thank you for the chats, both appropriate and inappropriate.

Finally, I would like to thank my friends outside of University and family for putting up with me these past four years! I'm sure it's been difficult at times. Rachel Salem, my oldest and best friend, thank you for the hugs, laughs, memories and for always being there for me! Also thanks to Ellie Bowen, your dry sense of humour and love of shopping and partying has got me through! Tom and Katie Harwood, thank you for being lovely people, I am so lucky to have you both as friends.

Mum and Dad, thank you for being such wonderful parents, I could never have asked for anyone better. You two are the best and most kind and loving people I know. You are perfect and I am so proud to call you my parents. I also want to thank my brother Adam for feeding me and keeping me entertained during my three-day time courses. And thank you to my younger brother Matthew, no matter how stressed I am, a cuddle and a play with you makes it all better.

Last but by no means least, thank you to my favourite person Mark Trueman, for all the laughs and love over these last few years. You are amazing and have kept me sane.

Chapter 1 – Introduction

1.1. Relevance of CAM in an increasingly arid world

Research into (Crassulacean Acid Metabolism) CAM plants is of rapidly increasing importance due to climate change and the associated challenges of global food and energy security. It has been predicted that the Earth's human population will surpass 9 billion people by 2050 (Parry *et al.,* 2007), and this is predicted to create significant further pressure on global agricultural productivity. The global average temperature has already risen by 0.8˚C since 1980 (Carlowicz 2010), and some areas of the globe have experienced even greater rises in average temperatures (Fig. 1.1). In 2007 the Intergovernmental Panel for Climate Change (IPCC) predicted a further 0.4˚C rise within the next two decades, and the scientific consensus suggests that this continuing warming is a direct result of the burning of fossil fuel reserves by humans, which releases CO_2 into the Earth's atmosphere that acts as a greenhouse gas, promoting the warming of the atmosphere. At present, approximately 40% of the world's land is considered either arid or semi-arid, preventing the efficient and productive cultivation of C_3 crops, which tend to do best in temperate climates with regular rainfall due to their relatively low water use efficiency (WUE). With average temperatures rising, and severe weather events including droughts becoming increasingly common, the global area of semi-arid and arid land is predicted to grow further by the end of the century placing even greater pressure on the world's food production systems.

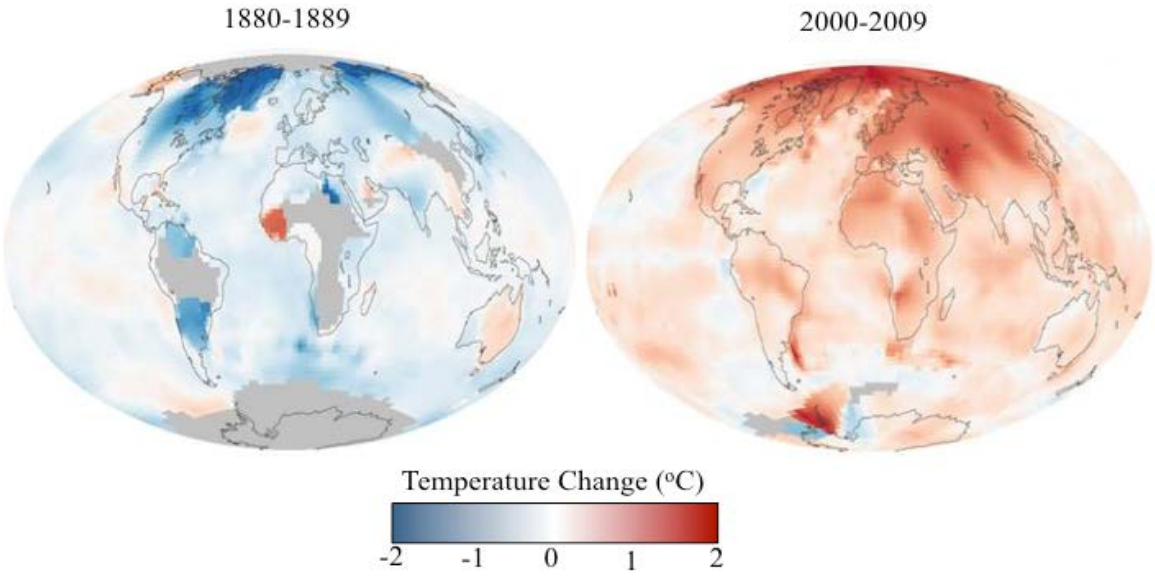

Figure 1.1. Temperature changes across the globe, reported by the IPCC and NASA (Images taken from Carlowicz, 2010). Grey patches indicate no data is available for those areas. Temperature change is expressed relative to average base temperature recorded in each area between 1951 and 1980. Overall it is easy to see the trend of temperature increase, making it increasing difficult for plants that are not adapted to coping with more arid climates to survive and grow productively, to feed a rapidly growing global population, and also meet the energy needs for this population, by bioenergy production in marginal lands.

The IPCC have also noted an increase in average minimum temperatures throughout the globe since 1960. Such climate change associated factors are likely to make it increasingly difficult to cultivate conventional C_3 and C_4 crops in the future, especially in the areas that have had rises in average temperature of multiple degrees, which are continuing to display dramatic increases in average temperatures (Fig. 1.1). These climatic changes are therefore likely to result in increasing the frequency and severity of extreme weather conditions such as heat and drought.

Furthermore there are also rising levels of atmospheric CO_2, which is likely to affect crop growth in the future. In the short-term, elevated CO_2 levels increase photosynthesis and growth is enhanced in most plant species (the so-called "CO_2-fertilisation effect") (De Souza

et al., 2008). However, as increasing atmospheric CO_2 also causes global warming, it has been predicted that there will in the future be a tipping point where the yield gains due to the CO_2-fertilisation effect are more than offset by the yield losses caused by elevated temperatures and the expected increased frequency of flood and drought events. In addition, long-term exposure of plants to high CO_2 can lead to elevated starch and sugar accumulation in leaves, which can induce feedback inhibition of photosynthesis (Sage *et al*., 1989; Stitt 1991). In some CAM plants, however, long-term exposure to elevated CO_2 does not have this detrimental effect on photosynthesis, suggesting that CAM plants could be the most ideal crop to grow in high CO_2, as well as high temperatures and the increased likelihood of severe drought areas (Nobel & Israel, 1994; Ceulemans *et al*., 1995).

CAM plants have also been shown to survive, with near maximal productivity, in degraded land with poor soil conditions, which is why many species are also of interest for sustainable biofeedstocks. CAM plants require 80 % less water compared to some C_3 and C_4 species, but produce similar amounts of biomass (Nobel 1996b; Borland *et al.,* 2009b; Somerville *et al.,* 2010; Borland *et al.,* 2011).

These global challenges are key factors underpinning the development of a number of different ambitious projects which aim to enhance the efficiency of crop photosynthesis as a way to improve agricultural productivity and viability (Covshoff *et al.,* 2012; Ducat *et al.,* 2012; Maurino *et al.,* 2013). Currently active projects include those that aim to engineer C_4 photosynthesis into C_3 plants (Covshoff *et al.,* 2012; Leegood 2013), carbon-concentrating mechanisms (CCMs) from cyanobacteria into C_3 crops (Price *et al.,* 2013; Zarzycki *et al.,* 2013), and CAM into C_3 plants (Yang *et al.,* 2015). All of these projects are seeking to use the latest plant synthetic biology approaches as their means to introduce the many 10s of genes involved in these photosynthetic adaptations into their target crop species (Borland *et al*., 2014; De Paoli *et al*., 2014; Yang *et al*., 2015).

1.2. Different photosynthetic mechanisms for carbon dioxide fixation – C_3, C_4 and Crassulacean acid metabolism (CAM)

CAM is one of three main classes of photosynthetic carbon dioxide fixation found in higher plants, the other two being C_3 and C_4. The three photosynthetic types are widely distributed throughout the globe, but in general, C_4 plants prefer warm, monsoonal climates, whilst CAM plants tend to be more common in water-limited habitats, and C_3 predominates in temperate regions.

C_3 is the ancestral photosynthetic pathway, which is utilised by the majority of the world's plants. C_3 plants use the enzyme ribulose 1,5-bisphosphate carboxylase oxygenase (RuBisCO) in the Calvin cycle in their chloroplasts to fix atmospheric CO_2 into sugars for growth. Whilst the carboxylase activity of RuBisCO is essential for carbon fixation into sugars, the oxygenase activity is viewed as an evolutionary relic that leads to a wasteful side reaction known as photorespiration. When O_2 is incorporated into ribulose 1,5-bisphosphate by RuBisCO, the product, 2-phosphoglycolate, must be recycled in order to maintain a balanced Calvin cycle, which has sufficient ribulose bisphosphate for CO_2 fixation to continue. Phosphoglycolate is recycled through a series of biochemical steps referred to as photorespiration, which lead to a loss of CO_2 and NH_3, and the consumption of ATP. Photorespiration has been estimated to reduce the efficiency of C_3 photosynthesis by up to 40% (Ehleringer and Monson, 1993).

At temperatures around and above 25˚C and atmospheric CO_2 concentrations approximately 390 μmol $mol^{-1}$and below, RuBisCO's oxygenase function occurs at high levels, despite appearing to have no useful function. This results in high photorespiration levels (Ehleringer & Monson, 1993). C_3 plants thus perform less well under sub-tropical and tropical conditions where high temperatures are common all year round.

To overcome the oxygenase side-reaction and photorespiration, C_4 and CAM plants have evolved carbon-concentrating mechanisms that pump CO_2 around RuBisCO, and thereby favour its carboxylase activity and outcompete its oxygenase activity. Plants conducting C_4 have been shown to have twice the photosynthetic efficiency of C_3 species in hot environments (Zhu *et al*., 2008). C_4 photosynthesis has increased efficiency due to spatial separation of primary and secondary CO_2 fixation, where Phosphoenolpyruvate carboxylase (PEPC) fixes CO_2 as a 4-carbon organic acid in the mesophyll cells, and CO_2 is then decarboxylated in the bundle sheath, which concentrates CO_2 around RuBisCO, thus repressing its oxygenase function (Sage *et al.,* 2012). CAM on the other hand separates the CO_2 fixation events temporally, where primary fixation occurs at night, and re-fixation occurs behind closed stomata during the day (Griffiths *et al.,* 2002).

In C_3 and C_4 photosynthesis, net CO_2 assimilation occurs in the day when air temperatures are higher and humidity is lower. This means that the loss of water during transpiration is relatively high for these species. In CAM plants, CO_2 assimilation occurs at night when air temperatures are lower and humidity is higher. This helps to minimise transpirational water loss and can lead to water use efficiency (WUE) that is up to 6-fold greater than C_3 plants (Nobel 1996a; Borland *et al.,* 2009a).

It has been estimated that CAM is occurs in around 16000 species, which are spread across 343 genera and at least 36 families; totalling approximately 6 - 7% of all vascular plants (Winter *et al.,* 1996a; Crayn *et al.,* 2004). This photosynthetic adaptation is thought to have evolved from the C_3 ancestral state on multiple independent occasions in many of these lineages. In several cases, there is good evidence for independent origins of CAM within individual families, especially for Orchidaceae and Bromeliaceae (Crayn *et al.,* 2004; Silvera *et al.,* 2009).

1.3. Crassulacean Acid Metabolism (CAM)

CAM is a metabolic adaptation of photosynthetic CO_2 fixation that concentrates CO_2 around RuBisCO via a temporally regulated biochemical pathway involving separate primary and secondary CO_2 fixation steps. By performing initial CO_2 fixation in the dark into the 4-carbon organic acid, malic acid, CAM plants are able to take advantage of reduced rates of evapotranspirational water loss that occur at night, and thus are able to achieve much higher WUE than plants that perform CO_2 fixation in the light when it is hotter and drier. CAM plants are currently of particular interest due to the fact that they can grow productively in semi-arid regions, where C_3 and C_4 crops would struggle to achieve a respectable yield. Several model species are currently under investigation as exemplar systems that will allow the elucidation of the key components of CAM and its daily temporal regulation. Amongst these, *Kalanchoë fedtschenkoi* has in recent years become one of the most useful CAM models due to the wealth of earlier CAM biochemical and physiological discoveries that were achieved in this species (Hartwell, 2006), combined with the more recent development of genome and transcriptome databases, and an efficient stable transformation protocol (Dever *et al*., 2015, Yang *et al*., 2015). These resources make *K. fedtschenkoi* currently the most tractable model system for the dissection of all of the components of CAM, from the genes and proteins involved in the core biochemistry, to those involved in CAM's daily regulation, its developmental and stress-responsive control.

CAM and the high WUE it bestows are key aspects of *Kalanchoë's* ability to grow successfully in the seasonally dry environment under which it evolved in Madagascar. CAM is characterised by inverse stomatal behaviour, whereby stomata open for atmospheric CO_2 uptake and fixation during the cooler, more humid night, and close throughout the hottest, driest part of the light period, thereby preventing water loss. CO_2 supply for photosynthesis is maintained behind closed stomata during the light period through the decarboxylation of the

stored malate, which was accumulated through primary atmospheric CO_2 fixation during the previous dark period.

1.3.1. Phases of CAM

The CAM pathway occurs within individual leaf mesophyll cells where this biochemical adaptation is controlled temporally over the 24 h light/ dark cycle. The temporal pattern of CAM has been conveniently segmented into four key phases based on the classic physiological measurements of CO_2 fixation (Fig. 1.2; Osmond, 1978, Borland 2009). These phases emphasise the temporal separation of both primary and secondary CO_2 fixation between the light and dark periods.

Phase I commences at dusk. Nocturnal stomatal opening occurs, and CO_2 is taken up by the leaf and fixed directly from the atmosphere via PEPC, forming oxaloacetate (OAA). OAA is converted to malate by malate dehydrogenase (MDH), and stored as malic acid in the vacuole, where its levels accumulate throughout the dark period.

At the beginning of the photoperiod, phase II occurs. Stomata have begun to close, but are still sufficiently open for both PEPC and the light-activated RuBisCO, to fix CO_2 for a short period. This results in a spike in CO_2 fixation (Fig. 1.2), and both C_3 and C_4 products can be observed (Griffiths *et al*., 1990).

During Phase III the stomata are closed coincident with the main period of decarboxylation of the malic acid from the previous night. There is no net CO_2 assimilation during this time, but inside the leaf there is a high internal partial pressure of CO_2, and CO_2 is re-fixed by RuBisCO via the Calvin cycle into sugars.

Phase IV of CAM occurs in the latter part of the light period when the stores of malate have been exhausted and the internal partial pressure of CO_2 declines. In this phase, the stomata reopen and conduct normal C_3 photosynthesis via direct fixation of atmospheric CO_2 into the Calvin cycle for the remainder of the photoperiod (Borland *et al*., 2009). It must be noted though that environmental conditions can alter or prevent this phase from occurring. For example, water deficiency will abolish phase IV (Cushman, 2001).

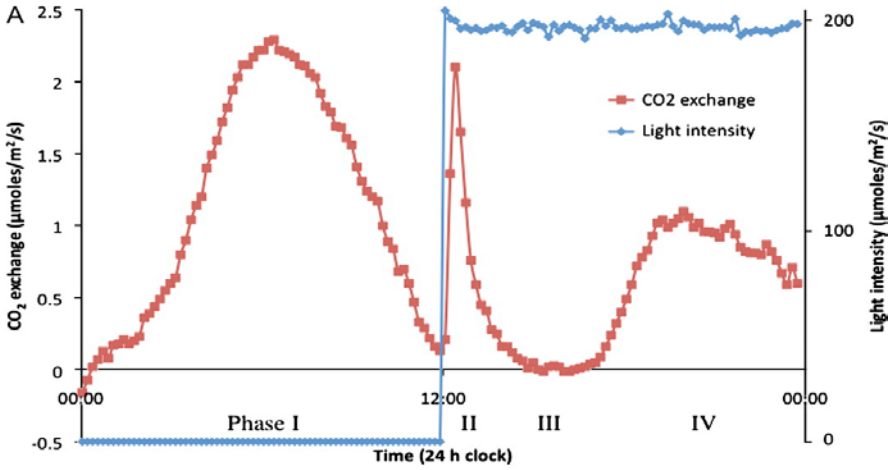

Figure 1.2. CO_2 exchange over a 24 h CAM cycle in *Kalanchoë fedtschenkoi.* In phase I the stomata open allowing PEPc to fix atmospheric CO_2. In phase II the stomata are still open and CO_2 can be fixed by both PEPc and RuBisCO for a short while. The stomata are closed in phase III, and Rubisco re-fixes CO_2 from the decarboxylated malate. If the plant is well watered and malate stores are exhausted, then the stomata will re-open and RuBisCO fixes atmospheric CO_2 directly in phase IV (Borland *et al*., 2009).

1.3.2 Biochemistry of CAM

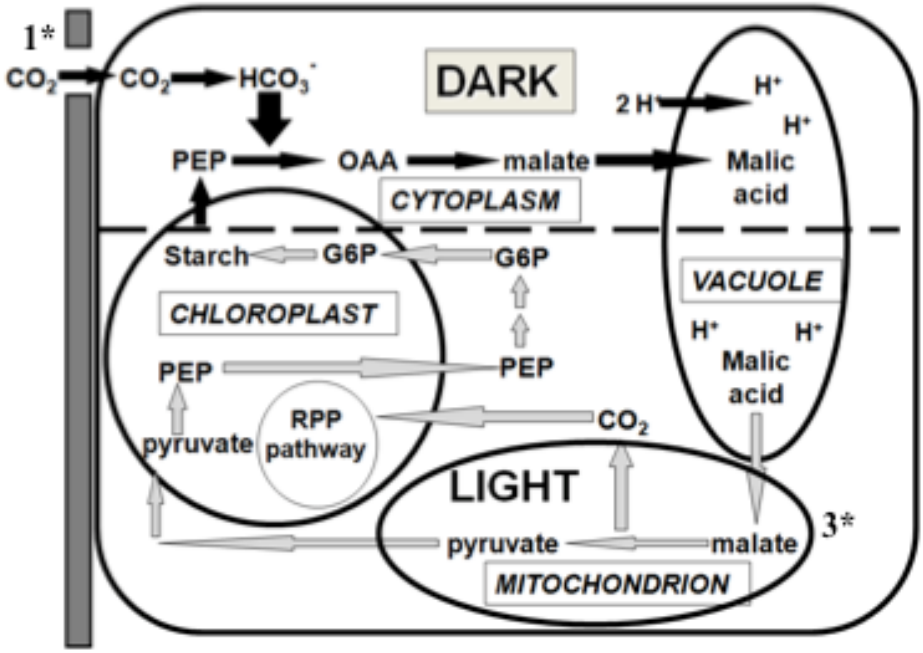

Figure 1.3. Overview of CAM. Black arrows represent dark stages and grey arrows represent light stages (Borland *et al.,* 2009b). CO_2 enters the leaf when stomata are open at night. CO_2 is fixed by phosphoenolpyruvate carboxylase (PEPC), which produces oxaloacetate. Malate dehydrogenase then converts oxaloacetate to malate, which is transported into the vacuole, and stored as malic acid. In the light, malate is transported out of the vacuole and is broken down to CO_2 and pyruvate by a decarboxylase enzyme. Pyruvate is recycled to form storage carbohydrates, and CO_2 is fed into the Calvin cycle and refixed by RuBisCO behind closed stomata.

The outline of the four-phase framework for the temporal pattern of CAM described above provides a useful overview of key physiological aspects of the pathway, but does not provide a detailed description of the biochemical steps and their regulation. A full appreciation of the biochemistry and its regulation are important for understanding the detail of the daily rhythm of CAM.

At the beginning of the dark period, when CAM plants open their stomata (1*, Fig 1.3), starch breakdown has commenced feeding sugars into glycolysis, which provides the phosphoenolpyruvate (PEP) required for PEPC. A cytosolic carbonic anhydrase converts

atmospheric CO_2 to bicarbonate (HCO_3^-). PEPC uses PEP as the acceptor substrate to fix the HCO_3^- to OAA. Malate dehydrogenase converts the OAA to malate, which is then transported into the vacuole via a voltage-gated inward-rectifying anion channel. Inside the vacuole malate is stored as malic acid due to the addition on two hydrogen ions (H^+). This occurs as there is a high concentration of H^+ ions inside the vacuole. The high proton concentration inside the vacuole is generated by the tonoplast membrane H^+-ATPase and H^+-PPase (Martinoia *et al.,* 2006). Malic acid carries on accumulating throughout the dark period, and can reach concentrations of up to 200mM at dawn (Borland *et al*., 2009). The amount that accumulates is dependent on many factors, such as vacuolar storage capacity and the availability of storage carbohydrate reserves (Griffiths, 1992).

In the light period, malate is exported from the vacuole to the cytosol, possibly via the tonoplast dicarboxylate transporter (Borland *et al*., 2009)(3*; Fig 1.3). Different CAM species use different decarboxylase enzymes to decarboxylate malate in the light (Dittrich, 1976; Christopher and Holtum, 1996). The three enzymes used are $NADP^+$-malic enzyme (ME), NAD^+-ME, or a combination of malate dehydrogenase plus PEP carboxykinase (Christopher and Holtum, 1996). Malate decarboxylation generates high internal CO_2 concentrations inside the leaf mesophyll cells. The Calvin cycle enzyme RuBisCO is activated in the light via RuBisCO activase, and re-fixes the CO_2 from malate decarboxylation via its carboxylase activity. Sugars are produced via the Calvin cycle, yielding sucrose for export and growth, plus starch accumulation in the leaf to fuel nocturnal PEP provision and respiration (Cushman, 2001). In ME-type CAM species, pyruvate orthophosphate dikinase (PPDK) converts the pyruvate produced as a product of malate decarboxylation to PEP, which then enters gluconeogenesis and is recycled to starch via ADP-glucose pyrophosphorylase. PPDK has been reported in both the chloroplast and/ or cytosol in different CAM species suggesting that some species must transport pyruvate into

the chloroplast for conversion to PEP, whilst some species are able to convert pyruvate to PEP in the cytosol.

Malate decarboxylation generates a high internal partial pressure of CO_2 inside the leaf, which is believed to signal stomatal closure in the light. This leads to the concentration of CO_2 around RuBisCO in the chloroplast, favouring the enzyme's carboxylase activity over its oxygenase acivity. This minimises the wasteful side reaction of photosynthesis known as photorespiration, which in C_3 plants recycles phosphoglycolate from the oxygenase activity of RuBisCO leading to a loss of CO_2 and ammonia (NH_3), and consumption of adenosine triphosphate (ATP).

1.3.3. The evolutionary continuum of CAM species

It has been suggested that CAM has evolved independently many times and is found in at least 36 families, representing about some 6 - 7% of all higher plants. It is therefore perhaps not surprising that there are numerous variations in CO_2 assimilation, acid flux, carbohydrate sources for nocturnal PEP, and stomatal behaviour amongst the diverse species that use CAM (Winter *et al.,* 1996b). As well as different species showing differences in these key facets of CAM, the environment can also impact on these characteristics. The weakest form of CAM is known as CAM-cycling, where daytime net CO_2 uptake occurs, and some respiratory CO_2 is refixed at night, producing small C_4 acid fluctuations (Luttge, 2004). There are also facultative CAM species, which use the C_3 pathway to maximise growth when water is abundant, but later undergo a C_3-CAM transition, which is often linked to moisture availability or other environmental stressors, with the best example of this being the common ice plant, *Mesembryanthemum crystallinum* (Chu *et al.,* 1990). Obligate/constitutive CAM species are plants that show almost exclusive night time (phase I) atmospheric CO_2 fixation,

and also show large malic acid fluctuations, such as *Opuntia ficus-indica* (Acevedo *et al*., 1983). CAM-idling is referred to as the 'strongest' type of CAM, but only normally occurs when CAM plants are undergoing severe drought. This is where stomata are closed throughout the light/ dark cycle, preventing water-loss through open stomata giving these plants greatly increased WUE. Plants conducting CAM-idling re-fix respired CO_2, and so still conduct diel cycling of organic acids, but at a much lower magnitude (Luttge, 2004).

In some CAM species, for example *K. fedtschenkoi*, not all of the leaves are 'obligate' for CAM (Jones 1975; Sipes *et al.,* 1985). There is a developmental progression into CAM as leaves age down the stem from the youngest leaves performing C_3 at the shoot apical meristem (SAM) to fully developed leaves performing full CAM further down the stem (Jones, 1975).

1.3.4. CAM development in *K. fedtschenkoi*

As already emphasised above, *K. fedtschenkoi* has recently become one of the favoured model systems for functional genomics research on CAM. It belongs to the family Crassulaceae, in which CAM was first discovered, and is endemic to Madagascar. Although this species performs CAM in its older leaves even when well-watered, there is a clear developmental progression from C_3 to CAM as leaves emerge from the shoot apical meristem (SAM) and mature (Fig 1.4A). Leaf pair one (LP1; youngest leaves either side of the SAM) perform C_3 photosynthesis; capturing atmospheric CO_2 through open stomata in the light (Figure 1.4A & 1.4B). LP2 and older all perform some degree of nocturnal stomatal opening and CO_2 fixation indicative of CAM. For example, the gas exchange data for LP3 in Fig. 1.4B reveal that these leaves perform all four phases of CAM, but that the magnitude of their nocturnal CO_2 fixation is relatively low compared to LP6, which perform full CAM (Fig.

1.4B). LP6 and below perform the majority of their CO_2 fixation in the dark with negligible light-period stomatal opening or direct fixation of atmospheric CO_2 (Fig. 1.4B; Kluge and Ting, 1978; Osmond, 1978, 2007; Nobel, 1988; Winter *et al*., 2008). This developmental profile of CAM has proved to be an extremely powerful experimental tool for comparative transcriptional profiling of CAM development and light/ dark regulation in *K. fedtschenkoi*. Over the last five years, the Hartwell lab have undertaken a comprehensive RNA-seq analysis of the transcript changes associated with CAM induction and its light/dark and circadian control in *K. fedtschenkoi* by comparing the transcriptome of LP1 and LP6 using a combination of *de novo* sequencing and assembly of the genome and transcriptome using a combination of 454 and Illumina sequencing technology, and quantitative analysis of the regulation of the transcriptome between different leaf pairs sampled at intervals over the light/ dark cycle using RNA-seq with the Applied Biosystems SOLiD and Illumina HiSeq sequencing systems (further details are in Chapter 3).

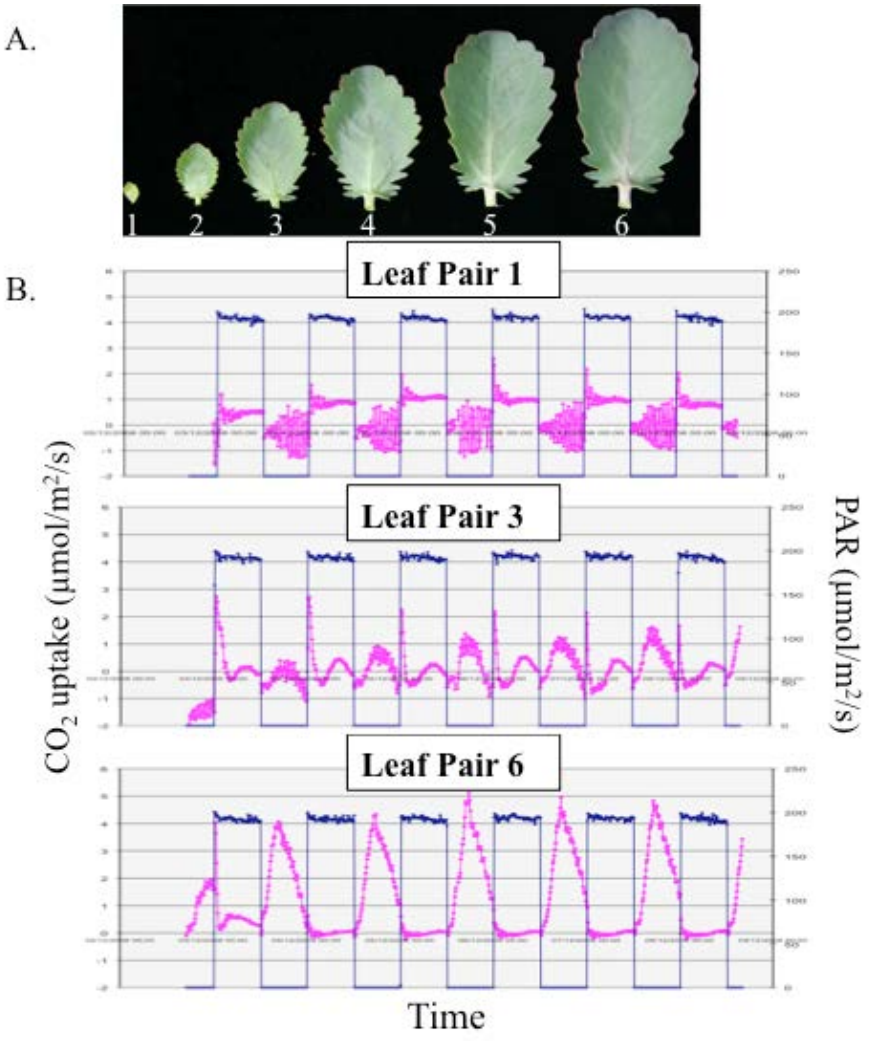

Figure 1.4. *Kalanchoë fedtschenkoi* leaf pairs 1 to 6, and the associated CO_2 fixation graphs for leaf pairs 1 and 6. **A.** Images of a leaf taken sampled from the 6 leaf pairs. The corresponding leaf pair number is under the images. **B.** Gas exchange data collected from leaf pair 1 (top) leaf pair 3 (middle) and leaf pair 6 (bottom) in 12:12 L:D, showing light intensity (blue) and CO_2 fixation (pink). In leaf pair one, CO_2 is fixed during the light period, whereas by leaf pair 3 fixation is also beginning to start at night, and by leaf pair 6, CO_2 is now carried out during the dark period. Gas exchange data collected by Boxall *et al.,* (unpublished data).

1.4. Temporal Regulation of CAM by the Circadian Clock

1.4.1. The Circadian Clock

The circadian clock provides an internal estimate of time, which can synchronise biological events with the Earth's 24 h daily rotation on its axis (Fig. 1.5). These rhythms are able to persist under constant environmental conditions, and are temperature compensated, but can be reset by certain stimuli, such as extreme heat and light (Harmer, 2009). In CAM plants, it has even been demonstrated that the clock is buffered against abiotic stresses too (Boxall *et al.,* 2005). These rhythms have even been found in etiolated seedlings, which have never been exposed to either a temperature step or light treatment, which suggests the plant circadian clock is endogenous (Salome *et al*., 2008; Harmer, 2009). It has been shown in plants that these rhythms display cell autonomy, and it has also been shown that there is little/no intercellular interaction between cells (Harmer, 2009; Thain *et al.,* 2002, Rascher *et al.,* 2001; Thain *et al*., 2000). CAM is believed to be coordinated relative to the light/dark cycle by the plant circadian oscillator. Despite the best characterised plant clock having been deciphered in Arabidopsis, many of the findings are applicable to most other angiosperms. Research has discovered that there is a considerable level of evolutionary conservation between clock components, their function and architecture across diverse members of the plant kingdom (Song *et al.,* 2010). Furthermore, in many cases the clock genes cycle with phase that matches that of their Arabidopsis counterparts. It is interesting that the smallest free-living photosynthetic eukaryote: *Ostreococcus tauri*, has only been shown to possess homologs to *CCA*1 and *TOC1*, which repress one another. This suggests that a two component core oscillator may be sufficient for such single-celled organisms, but as multicellular photosynthetic eukaryotes evolved, it seems that the core plant clock developed increasing levels of complexity, presumably in order to regulate the higher temporal

complexities required for the success of a multi-cellular, multi-organ photosynthetic eukaryote.

It has been shown through use of mathematical modelling that interlocking feedback loops of interacting genes and their protein products provide the clock with robustness, preventing perturbation, suggesting that more of the clock pathways and feedback loops are yet to be discovered (Kwon & Cho, 2007; Tsai *et al.,* 2008).

The best and most up-to-date model for the core plant circadian clock involves a number of genes, most of which encode transcription factors or proteins involved in targeted protein degradation. *LATE ELONGATED HYPOCOTYL* (*LHY*) and *CIRCADIAN CLOCK ASSOCIATED 1 (CCA1)* are dawn-phased single-MYB repeat transcription factors. *TIMING OF CAB EXPRESSION 1 (TOC1*) reciprocates with *CCA1/LHY* and is a pseudo-response regulator (PRR) protein that has been demonstrated to act directly as a transcription factor (Gendron et al., 2012). When CCA1/ LHY bind to the *TOC1* promoter at the evening element (AAAATATCT), they repress its expression for the majority of the light period. As CCA1/ LHY decline in the latter half of the light period, *TOC1* repression is relieved leading to the observed rise in *TOC1* transcript levels at dusk. TOC1 itself has been demonstrated to act as a transcriptional repressor, directly binding DNA through its CCT-domain. TOC1 binds to a motif within the *CCA1/ LHY* promoters, which has the sequence TGTG at its core, and represses their expression (Gendron et al., 2012). Further components of the current Arabidopsis clock model include *PRR7*,*PRR9*, *GIGANTEA (GI),* and *ZEITLUPE (ZTL).* Orthologs of all of these core clock genes have been identified in several CAM species, including *M. crystallinum* and *K. fedtschenkoi* (Boxall *et al*., 2005; Dever *et al*., 2015)

The clock therefore consists of interlocking negative feedback loops (Fig. 1.5). At dawn transcript levels of *CCA1* and *LHY* peak due to induction by GI (Fig 1.5). CCA1/LHY

proteins act to repress the evening gene *TOC1* and also feedback to *GI*, causing inhibition of transcription. CCA1/LHY also promote the expression of *PRR7* and *PRR9*, which are day-time genes, which in turn feedback and repress *CCA1* and *LHY*. When CCA1 and LHY protein levels are low this then enables *GI* and *TOC1* to be expressed, and TOC1 leads to further repression of *CCA1* and *LHY* (Gendron *et al.,* 2012). GI also induces further *TOC1* expression, which TOC1 then represses the expression of. TOC1 is expressed throughout the night until the protein is targeted for proteasome-dependent degradation by ZTL. For efficient degradation a ZTL-GI complex must form for stability, and this allows the morning genes *CCA1* and *LHY* to be transcribed due to TOC1 degradation removing their transcriptional repression, enabling transcript for the morning genes to accumulate.

In addition to the core circadian clock, which has been found to be highly conserved between most angiosperms, there are also many other oscillator components interconnecting with the core to form multiple feedback loops and output pathways. These have been studied in some detail in Arabidopsis due to the wealth of genetic and transcriptome data available. From the confirmation that the core clock is conserved in angiosperms, including both Arabidopsis and the CAM plant *M. crystallinum*, then it is likely that many more oscillator components will also function in a similar way across and between species.

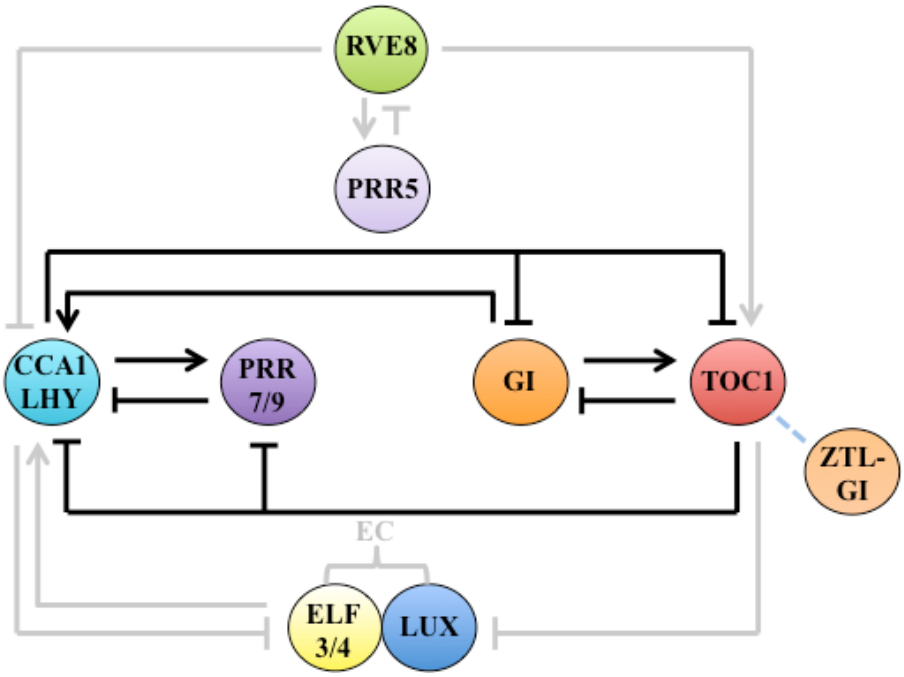

Figure 1.5. The plant circadian clock. The core clock involves CCA1/LHY, TOC1, PRR7/9, GI and ZTL, and interactions are shown in black lines. This clock has also been confirmed in CAM plant *M. crystallinum*. Other factors shown to play roles in the Arabidopsis circadian clock, which provide extra feedback loops, and so making the clock more robust are: RVE8, PRR5, ELF3/4 and LUX. This diagram shows a simplified version of the clock, with many more genes being implicated.

Other key regulators of the Arabidopsis clock are shown in Figure 1.5 and interactions with the core oscillator are shown in grey lines: *PSEUDO RESPONSE REGULATOR5 (PRR5), LUX ARRHYTHMO (LUX), EARLY FLOWERING3 (ELF3), EARLY FLOWERING 4 (ELF4), ZEITLUPE (ZTL) and REVEILLE4/6/8 (RVE4/6/8).*

PRR5 is induced by *RVE8*, which belongs to the same family of transcription factors as *CCA1/LHY*. Interestingly, in Arabidopsis RVE8 appears to mediate the opposite function to *CCA1/LHY*. CCA1/LHY repress *TOC1* expression, due to deacetylation of the *TOC1* H3 histone, whereas RVE8 promotes acetylation, and so promotes transcription (Farinas *et al.,* 2011). RVE8 also promotes PRR5 induction, which in turn feeds back to *RVE8* leading to inhibition. RVE8 and PRR5 also functions to inhibit *CCA1/LHY* expression. LUX, ELF3 and ELF4 proteins interact to form the evening complex (EC), and all three genes are repressed

by *CCA1/LHY,* but once this EC is present it is able to feed back and inhibit *CCA1/LHY* expression (Hsu *et al.,* 2013). PRR7/9 also function to repress *CCA1/LHY*.

1.4.2. Advantages of circadian clock and environment synchronisation.

It has been demonstrated that correct circadian timing bestows huge advantages on a plant that has a circadian clock that is synchronised with its external environment, known as 'circadian resonance' (Dodd *et al.,* 2005). Dodd *et al*. (2005) demonstrated that *A. thaliana* plants with their clocks tuned to their external environment contained more chlorophyll, fixed more carbon, were able to grow faster and were also able to survive better when there was competition between plants. Mutants in which the core circadian clock gene *CIRCADIAN CLOCK ASSOCIATED1* (*CCA1*) was constitutively over-expressed showed arrhythmia in CO_2 uptake, resulting in a 50 % reduction in photosynthetic CO_2 fixation (Dodd *et al*., 2005). Also, *ztl* mutants have a longer circadian free-running period length, for both gene expression and CO_2 assimilation (Wang *et al.,* 1998; Dodd *et al.,* 2004). It has also been demonstrated in rice that a *gi* suppressor mutant showed decreased chlorophyll, altered carbon metabolism and increased stomatal conductance, as well as a severe 'dwarf phenotype' (Izawa *et al.,* 2011).

Furthermore, Graf *et al*. showed that starch degradation was controlled by the circadian clock, and Arabidopsis mutants with *cca1/lhy* mutants showed inappropriate rates of starch degradation, which led to carbon starvation and had a negative impact on nocturnal metabolism and growth (2010).

1.4.3. The Circadian Clock is highly complex

The clock is often said to have three sections: input pathways, the core oscillator of positive-negative feedback loop that set the time, and output pathways (Harmer 2009). From the intensive research conducted on the plant circadian clock in Arabidopsis over the last 15 – 20 years, we now know this is a highly oversimplified view. The core circadian clock can perhaps better be considered as a signal integrator within a complex network. Inputs to the clock are able to regulate its function at multiple points, but the clock can also gate itself from inputs such as light by using the output pathway. Output pathways have also been shown to be able to regulate themselves (Harmer 2009). Therefore, a gene cannot be placed neatly within one of the three sections within the clock, due to it conducting multiple functions (Harmer 2009).

As well as TOC1 functioning within the core clock, it also displays a role in light signalling, a function that has also been demonstrated for other PRRs too (Mas *et al.,* 2003; Martin-Tryon *et al.,* 2007; Martin-Tryon *et al.,* 2008). Additionally, as well as ZTL regulating TOC1 levels in a light dependent manner, it is also required for clock persistence in constant darkness (Somers *et al.,* 1998; Devlin *et al.,* 2000; Kevei *et al.,* 2007). GI has been shown to play a role in light regulation, but also plays a biochemically separate role in regulation of flowering time by promoting expression of florigens *CONSTANS (CO)* and *FLOWERING LOCUS T (FT)* (Mizoguchi *et al.,* 2005; Martin-Tryon *et al.,* 2007). ELF3 and 4 are also required in constant conditions to enable the circadian clock to continue to robustly oscillate, by negatively regulating light input into the clock, especially in constant light conditions (Mcwatters *et al.,* 2000; Covington *et al.,* 2001; Ding *et al.,* 2007; Mcwatters *et al.,* 2007). The evening complex (LUX, ELF3&4) has also been shown to be involved in hypocotyl growth by inducing transcription of *PHYTOCHROME-INTERACTING FACTOR 4* and *5 (PIF4&5)* during the dark (Nusinow *et al.,* 2011).

As well as light, temperature also acts as an input to the plant clock, with transcription of many clock components being temperature sensitive, including *CCA1, LHY, TOC1* and *GI* (Salome *et al*., 2005; Gould *et al*., 2006; Paltiel *et al*., 2006; Salome *et al*., 2006). Furthermore PRR7 and 9 are key proteins for the response of the clock to temperature changes, as a *prr7 prr9* double mutant could not be entrained to temperature cycles, and also did not respond to temperature pulses (Salome *et al.,* 2005).

As well as affecting the transcription level, many clock genes can also have different sequences translated due to alternative splicing (AS). In Arabidopsis 61% of intron-containing genes show AS, and this is only when assayed under normal growth conditions at one developmental stage (Marquez *et al.,* 2012). Intron retention is the most common form of AS, whilst exon skipping events also occur, which can lead to unproductive mRNA, or effect stability or transcript levels (Filichkin *et al.,* 2010). In 2010, Filichkin *et al*., demonstrated that due to varying light and temperature conditions, *CCA1* would undergo intron 4 retention. Then in 2012, James *et al*. demonstrated that when temperature conditions were manipulated, 15 key AS events were seen in 7 key clock genes: *LHY, CCA1, TOC1* and *PRR*s *3, 5, 7* & *9*. This led to changes of abundance and duration of transcript presence for these clock genes, which shows an added layer of clock regulation complexity (James *et al.,* 2012).

1.4.4. Crosstalk between the clock and metabolism

As well as core circadian clock genes having multiple roles to play within the clock itself and through linking to output pathways, plant metabolism and the clock also interact with one another (Sanchez *et al.,* 2011; Eckel-Mahan *et al.,* 2013; Haydon *et al.,* 2013). Not only does the clock regulate metabolism, but metabolism can also feedback to influence clock functions.

Circadian regulated phytohormones have been shown to be able to feed back and module the plant clock, which is already a well-known integration system in mammals (Reppert 2000). Phytohormones is the collective term for plant hormones such as: cytokinins, auxins, brassinosteriods (BR), ascisic acid (ABA), gibberellins (GA), ethylene and salicyclic acid (SA). BR, ABA, GA, auxin and cytokinin have all been shown to have their levels affected by light signals. Hanano *et al*., (2006) have shown that these phytohormones can feedback to influence the clock. Cytokinins have been shown to delay circadian phase, auxins play roles in regulating clock amplitude and precision, and BR and ABA control circadian periodicity (Hanano *et al.,* 2006; Salome *et al.,* 2006; Covington *et al.,* 2007).

Soluble sugars are another group of metabolites that have been shown in Arabidopsis to modify the expression of up to half of the known clock-regulated genes (Blasing *et al.,* 2005). This was shown using phosphoglucomutase (*pgm)* mutants, which lacked the ability to synthesise leaf storage starch, and so also lacked the buffering of starch turnover for the clock. This resulted in much larger changes in sugar levels over the 24 h cycle, and also amplified transcription on thousands of genes (Blasing *et al.,* 2005). Haydon *et al*., also showed cross talk between sugar metabolism and the clock, by showing rhythmic sugar signals can entrain the circadian clock and signal a ‘metabolic dawn’ through PRR7 (2013).

Nitrogen assimilation has been demonstrated to be regulated by the core clock gene *CCA1*, with overexpression lines showing changes in key nitrogen metabolism genes, but nitrogen levels have also been shown to feed back to CCA1 and modulate its phase of expression (Gutierrez *et al.,* 2008). Glutamate (Glu) and glutamine (Gln) appear to be able to feedback to the clock to indicate the status of nitrogen metabolism, with pulses of Glu leading to advances in *CCA1* phase, and Gln leading to delays in *CCA1* expression and also a shortened period (Gutierrez *et al.,* 2008).

Cytosolic signalling molecules such as cyclic adenosine diphosphate ribose (cADPR) also regulate the abundance of clock gene transcripts (Dodd *et al.,* 2007). cADPR drives Ca^{2+} oscillations by regulating vacuolar channels that control the release of Ca^{2+} into the cytosol. It is thought that changes in the level of cADPR are able to signal changes in physiology to the clock, enabling it to adapt to daily environmental changes.

Each of these feedback loops between key metabolic pools and the core components of the clock improves the clock's robustness, enabling the plant to react dynamically to constantly changing conditions, and modulate responses according to the severity of the environmental or cellular perturbation. This supports the suggestion that more clock controlled pathways and feedback loops are yet to be discovered (Kwon & Cho, 2007; Tsai *et al.,* 2008).

1.4.5. Temporal Control of CAM

For CAM to occur efficiently, tight coupling between the circadian clock and CAM biochemistry must occur, to enable specific CAM genes to be activated at the correct times of day to prevent futile cycling and to maintain high water use efficiency (WUE).

Temporal and circadian control of CAM has been known for many decades, as the daily pattern of CO_2 fixation that is so characteristic of CAM was found to persist in constant free-running conditions as far back as the late 1950s (Wilkins 1959). Following this, CO_2 exchange was also shown to possess a high degree of temperature compensation, and robustly oscillates between 10 and 30^{o}C (Wilkins 1962; Anderson *et al.,* 1989b; Anderson *et al.,* 1989a). By feeding $^{14}CO_2$ to *K. fedtschenkoi* leaves, it was demonstrated that the product of the clock controlled nocturnal CO_2 fixation was malate (Warren *et al.,* 1961), which PEPC is subject to feedback inhibition from (O'leary 1982). PEPC activity and sensitivity levels change over the 24h period, with a low apparent K_i for malate during the light (0.3mM) and a higher K_i (3mM) during the dark (Nimmo *et al.,* 1987; Carter *et al.,* 1991; Borland *et al.,* 1999). This demonstrated that PEPC activity is under circadian control. Changes in sensitivity were shown to be due to reversible phosphorylation on a serine residue by PPCK (Nimmo *et al.,* 1984; Nimmo *et al.,* 1986; Nimmo *et al.,* 1987; Carter *et al.,* 1991). PPCK has since been shown to be a key circadian control point for CAM, as it is regulated at the mRNA level and diurnally oscillates in both light/dark and constant conditions (Hartwell *et al.,* 1996; Hartwell *et al.,* 1999). Despite this discovery though, nothing is yet known of this gene's regulators, or the signalling pathway underpinning this circadian control.

CAM occurs in each photosynthetic mesophyll cell of the leaf (or stem in cacti). Futile cycling of simultaneous CO_2 fixation to malate, and malate decarboxylation to release CO_2 in each mesophyll cell is prevented by strict temporal control of the primary and secondary CO_2

fixation steps associated with CAM, as otherwise this would lead to wasting large amounts of ATP and NAD(P)H.

Transgenic lines of *K. fedtschenkoi* over-expressing the iceplant *McTOC1* gene from the constitutively active cauliflower mosaic virus 35S promoter (CaMV35S) had CO_2 fixation rhythms that collapsed to arrhythmia under constant light and temperature free-running conditions (Dall'omo 2011). These *TOC1* over-expressing lines also showed a large reduction in vegetative yield relative to wild type plants, demonstrating that a loss of circadian control of CAM reduced yield by up to two-thirds (Dall'omo, 2011). These findings further emphasised the importance of correct circadian clock control within the CAM pathway. However, little is known about the mechanisms governing the circadian control of CAM. Robust circadian oscillations have been demonstrated for a plethora of genes in the facultative CAM species, *M. crystallinum* (Cushman *et al*., 2008), and more recently, unpublished quantitative RNA-seq data for CAM leaves of *K. fedtschenkoi*, along with semi-quantitative and quantitative real-time RT-PCR analyses have demonstrated robust light/ dark and circadian oscillations in the transcript abundance many of the genes associates with CAM. CAM plants are also known to display robust circadian oscillations in CO_2 exchange, along with coordinated fluctuations in CAM-associated metabolites such as malate, starch, and sugars.

RT-PCR was used in *M. crystallinum* to try and determine whether circadian control was exerted at multiple points of the CAM pathway, or just at the level of *PPCK* (Boxall and Hartwell, unpublished data). They found at least 25 genes which qualified as CAM-induced, circadian clock-controlled genes, and these genes peaked at various time points over the 24 h LD cycle (Hartwell 2005). Due to CAM evolving multiple times, it was important to confirm that this global control of CAM by the circadian clock was not species specific. Several other species including *K. fedtschenkoi* and several *Clusia* species have also been shown to use a

circadian clock controlled PPCK to achieve phosphorylation of PEPC in the dark for CAM (Hartwell *et al.,* 1999; Taybi *et al.,* 2000; Taybi *et al.,* 2004).

More recently, large-scale transcriptome sequencing projects have been undertaken for a range of CAM species including *K. fedtschenkoi, K. laxiflora, Agave sisalana*, *A. deserti*, *A. tequilana*, and *Talinum triangulare*. A key goal of these transcriptome sequencing projects has been to decipher which genes are involved in CAM and its optimisation and temporal coordination in response (Yang *et al*., 2015).

1.4.6. Key challenges in determining factors involved in the circadian control of CAM

An important area in both circadian clock and CAM research is identifying genes and proteins which are involved in the signal transduction pathway that links the circadian clock and the various steps of the biochemical pathway in CAM, including the regulation of *PPCK*. One of the main ways expression levels of genes can be regulated is by transcription factors (TFs). These proteins are able to bind directly to DNA, either alone or in complexes, which enables them to exert either an activator or repressor function onto the gene (Correa *et al.,* 2008a). TFs are therefore a key group, which seem to possess strong candidates for exerting circadian control on CAM genes.

1.5. Transcription Factors (TFs)

Growth and development of all organisms cannot occur without strict regulation of gene expression. The control of transcription initiation rates by TFs represents one of the most important means of modulation gene expression (Correa *et al.,* 2008b). TFs are grouped into different protein families due to their primary and/or three-dimensional structural similarities in conserved domains, such as DNA-binding and/ or multi-merisation domains.

Gene expression is mediated through sequence specific binding of these TFs to *cis*-acting elements, which are found in the promoter, enhancer and/or silencing regions for the corresponding gene, which often occurs in a tissue-specific, development-stage-specific or stimulus-dependent manner (Zhang 2003). This DNA-specific binding can then facilitate other components of transcriptional machinery to initiate mRNA synthesis, or can prevent transcription from occurring (Meshi *et al.,* 1995). Many of families of TFs are common to all eukaryotic lineages, but others have been shown to be unique to plants, such as the *DNA-BINDING WITH ONE FINGER* (*DOF*) family of transcription factors (Shigyo *et al.,* 2007).

Many transcription factors do not function alone, with many groups forming homo- and/or heterodimers within and between TF families, leading to a vastly increased variation of target sequences (Meshi *et al.,* 1995; Diaz *et al.,* 2002; Zou *et al.,* 2008). Many TFs also undergo post-translational modifications, which can also affect the activation state of the protein, or its affinity to its DNA-binding site (Dubos *et al.,* 2010).

1.5.1. Plant Transcription Factors

Large-scale genomic comparisons have shown that angiosperm TF lineages have undergone more intensive gene expansion when compared to animals and fungi (Meshi *et al.,* 1995). There are 11 major families of plant transcription factors, and many of these are plant specific, namely APETALA2/ETHYLENE-RESPONSIVE-ELEMENT BINDING PROTEIN (AP2/EREBP), AUXING RESPONSE FACTOR (ARF-AUX/IAA), DNA-BINDING WITH ONE FINGER (DOF), WRKY and NAC (Riechmann *et al.,* 2000). In addition, the following TF families are shared across many eukaryotes: BASIC HELIX-LOOP-HELIX/MYC (bHLH/MYC), BASIC LEUCINE ZIPPER DOMAIN (bZIP),

HOMEOBOX (HB), MADS, MYB and ZINC-FINGER PROTEIN OF THE C_2H_2 TYPE (Z-C_2H_2).

Table 1.1. Major families of plant transcription factors.

Gene family	Family exclusive to plants?	Gene family functions	Estimated TF number in *Arabidopsis*
AP2/EREBP	Y	Flower development, cell proliferation, secondary metabolism, abiotic and biotic stress responses.	150
ARF-Aux/IAA	Y	Auxin responses, development and floral meristem patterning.	42
bHLH/MYC	N	Anthocyanin biosynthesis, light response, flower development and abiotic stress.	100
bZIP	N	Seed-storage gene expression, photomorphogenesis, leaf and flower development, defence responses, ABA response and gibberellin biosynthesis.	100
DOF	Y	Seed germination, endosperm-specific expression and carbon metabolism	41
HB	N	Stem cell identify, cell differentiation, growth responses, anthocyanin accumulation and cell death of leaf, root and ovule.	90
MADS	N	Flower development, fruit development, flowering time and root development.	80
MYB	M	Secondary metabolism, cellular morphogenesis, abiotic and biotic stress response and circadian rhythms.	180
NAC	Y	Development, pattern formation and organ separation.	105
WRKY	Y	Defence response.	75
Z-C_2H_2	N	Flower development, flowering time seed Flower, seed and root development.	85

This shows the main functions in plants that the families play, the number estimated in *Arabidopsis,* and also which families have been shown to be exclusive to plants (highlighted in green). Arabidopsis information, and some gene family functions are taken from Riechmann &Ratcliffe (2000).

From Table 1.1 it can be seen that from the TFs putative functions, they regulate huge ranges of plant processes, and although we have identified that these families affect these roles in some way, there are still very few that have been genetically characterised, with their functions confirmed *in vivo/ in planta*.

1.6. Project Objectives.

The over-arching aim of this project was to define genes involved in the circadian control and optimisation of CAM and characterise their functions in detail. A pipeline enabling effective identification of transcription factors that could potentially be involved in regulating CAM genes needed to be developed, from using SOLiD sequencing data to identify potential TFs, to producing and screening transgenic lines for changes in CAM.

This project was involved in identifying lines with a functioning transgene, and then bringing forward a select few lines to conduct detailed molecular, biochemical and physiological testing on, to determine whether these TFs identified were indeed involved in the circadian optimisation of CAM.

Chapter 2 - Materials and Methods

2.1 Plant Material

2.1.1 Standard entrainment conditions

Kalanchoë fedtschenkoi 'Hamet et Perrier' plants were are grown in a compost mix (John Innes Potting Compost No. 3: Sinclair Compost: Perlite; at a ratio of 3:3:1) containing slow release nutrients (Osmocote) applied at the recommended level of manufacturer in 10.5 x 10.5 x 12 cm pots. Plants were propagated via either clonal stem cuttings or adventitious leaf margin plantlets from wild type stock plants that originated as the accession of *K. fedtschenkoi* that Prof. Malcolm Wilkins originally obtained from the Royal Botanic Gardens, Kew (Wilkins, 1959). All plants were grown initially in a greenhouse with supplementary lighting where the average conditions were 16 h light (06:00 to 22:00; ~250 µmoles m^{-2} s^{-1}; ~23˚C) and 8 h dark (22:00 to 06:00; 0 µmoles m^{-2} s^{-1}, 18˚C) cycles. Before harvesting leaves plants were moved to Snijders Microclima MC-1000 climate controlled plant growth cabinets (Snijders Scientific, Microclima) 12:12 LD cycle (400 µmoles m^{-2} s^{-1}: 0 µmoles m^{-2} s^{-1}, 25˚C:15˚C, 60%:80% humidity), Light:Dark cycles. Leaves were collected 2h before dusk and dawn in OX lines and 2h before dawn in RNAi lines, to test gene expression. Samples were collected 1h before light and dark to test malate and starch levels. Only leaf pair 6 (LP6) leaves were used for the experiments unless stated otherwise. All leaf samples for gene expression tests were frozen immediately in liquid nitrogen (L N_2) and stored at -80˚C.

2.1.2 Light/ dark time course experiments using CAM leaves

After 7-days of entrainment in plant growth chambers, three biological replicates of leaf pair 6 samples were collected every 4 h over a 12:12 LD cycle (400 µmoles m^{-2} s^{-1}: 0 µmoles m^{-2} s^{-1}, 25˚C:15˚C, 60%:80% humidity, Light:Dark cycles). The light period commenced at 00:00 h local time and ended at 12:00 local time, whilst the dark period ran from 12:00 to 24:00. Leaf sampling generally commenced at 02:00, with further samples collected at 06:00, 10:00, 14:00, 18:00 and 22:00. Leaf samples were wrapped in tin foil, labelled and snap frozen in liquid nitrogen before storage at -80˚C in a freezer until use.

For preliminary screens for the transcript level of the genes targeted by the RNAi and over-expression constructs, samples were collected 2 h before the end of light (14:00 h) and dark period (22:00 h) from plants that were entrained to 16:8 LD (400 µmoles m^{-2} s^{-1}: 0 µmoles m^{-2} s^{-1}, 25˚C:15˚C, 60%:80% humidity, Light:Dark cycles).

2.1.3 LL time courses

Plants were pre-entrained for a minimum of 7-days in the Snijders Microclima MC-1000 plant growth cabinets under 12:12 LD cycles (400 µmoles m^{-2} s^{-1} : 0 µmoles m^{-2} s^{-1}, 25˚C:15˚C, 60%:80% humidity, Light:Dark). A light and dark LP6 sample was collected in biological triplicate 2h before dusk and dawn (10:00 L and 22:00 D respectively). At the end of the dark period, the Snijders cabinet was reprogrammed to constant light, temperature and humidity conditions in order to place the plant under circadian free-running conditions, which test for underlying endogenous circadian control of the biological outputs that were measured (LL conditions: light fixed at 100 µmol m^{-2} s^{-1}, temperature fixed at 15˚C, and humidity fixed at 70% humidity). Three biological replicates of leaf pair 6 samples were collected every 4 h, starting 2 h into

LL for *KfCDF2_FL_14C*, or after 26h under LL conditions for all other LL time course experiments. Samples were collected every 4 h for 3 days and immediately frozen in liquid nitrogen before storage at -80˚C in a freezer.

2.1.4 Flowering time experiments

2.1.4.1 Trial – Three biological replicates of each plant, both wild type and each transgenic line, were grown under greenhouse conditions for 24 weeks until they were large enough to be tested for their flowering response. Younger *K. fedtschenkoi* plants lack the competence to respond to short days and flower. Plants were then transferred to Snijders Cabinets (Microclima), which were set to either 16:8 LD, 12:12 LD or 8:16 LD (400 μmoles m^{-2} s^{-1}: 0 μmoles m^{-2} s^{-1}, 25˚C:15˚C, 60%:80% humidity, Light:Dark cycles). Plants were photographed every week for 14 weeks to determine the length of time needed for flowering to occur for each line.

2.1.4.2 Experiment – Plants were grown for 24 weeks from developmentally synchronised leaf margins in 16:8 greenhouse conditions, and were then moved into a walk-in growth room. The first week plants were entrained to 16:8 LD (approximately 400 μmoles m^{-2} s^{-1}: 0 μmoles/m^{-2}/s^{-1},20˚C:15˚C, L:D) was sampled 2 h before dawn (22:00) and dusk (06:00). After that, timing was switched to 8:16 LD conditions to induce flowering, and the samples (described in detail below) were collected 2 h before dawn (22:00) and dusk (14:00). Both high intensity and sodium lights were used in a ratio 3:1. Samples were collected throughout the flower-induction and flower development process in 8:16 LD (Weeks 1,4,5,6,7&8) and were immediately frozen in liquid nitrogen and stored at -80˚C in a freezer. The tissue samples collected varied between weeks, as it depended what part of the flowering process the plants

had progressed to. The samples collected throughout the process were: SAM, stem, LP1, LP3, LP6, LP10, hooked stem, closed buds, open buds and flowers. Due to the size of SAM, LP1, buds and flowers, only one biological replicate was able to be collected for these, as each sample required the pooling of material from several individual plants in order to obtain sufficient tissue for the downstream RNA isolations. Error bars for the corresponding real-time quantitative RT-PCR experiments on a range of flowering genes therefore only represent technical replicates rather than biological replicates. Three biological replicates were collected for the other samples (LP3, LP6, LP10, stems and hooked stems) collected at each time point.

2.1.5 Creating transgenic lines

2.1.5.1 Generating sterile plantlets - *K.fedtschenkoi* – RNAi and FL overexpression lines

The binary construct design and production, and the resulting stable plant transformation for all of the *K. fedtschenkoi* transgenic lines described in this thesis was carried out by former research technician in the Hartwell Lab, Jana Knerova. Mature leaf pairs (leaf pair 6 and older) were removed from well-established wild type *K. fedtschenkoi* plants growing under greenhouse conditions and were washed in 70 % ethanol for 30 seconds. Surface sterilization was then conducted using 10% hypochlorite with 1% Tween 20 and leaves were stirred in a large beaker of this solution for 10 mins. Leaves were then rinsed 5 times with sterile distilled water, before being blotted dry on sterile filter paper; working within the sterile laminar flow bench. Leaf margins were then excised with a scalpel under sterile conditions and the

cut edges were placed in MS30 IIA TDZ media (Murashige and Skoog medium containing Gamborg B5 vitamins, 3% (w/v) sucrose, solidified with 8% (w/v) Phytoagar containing 1 mg l^{-1} thidiazuron (TDZ), and 0.2 mg l^{-1} indole acetic acid (IAA)). Plates were then sealed and left in growth rooms (16:8 light: dark, 100 µmoles/m^{-2}/s^{-1}, constant temperature of 20˚C) until sterile adventitious plantlets had developed on the leaf margins.

2.1.5.2 Generating sterile plantlets - *K.laxiflora* - HA Tagged overexpression lines

For the *K. laxiflora* stable transformations with the HA-tagged transcription factor over-expression constructs generated as part of this project, *K. laxiflora* seeds were sterilised in 70 % ethanol for approximately 3 minutes with constant agitation. Seeds were then washed twice in sterile distilled water, and susequently re-suspended in sterile 0.2 % Phytoagar. Seeds were then pipetted individually onto a ½ strength sterile MS30 plate (50 % Murashige and Skoog medium containing Gamborg B5 vitamins, 1% (w/v) sucrose, solidified with 8% (w/v) Phytoagar). Plates were sealed with Micropore tape and left to germinate in the walk-in growth room (16:8 light: dark, 100 µmoles m^{-2} s^{-1}, 20˚C) until sterile seedlings developed.

2.1.5.3 Primer design and PCR product synthesis

2.1.5.3.1 RNAi primers

Primer constructs were designed to be either completely or partially in the 3^1 UTR of the gene specific region (Table 2.1). The forward primer always begins with a CACC for directional insertion of the PCR fragments during pENTR/D TOPO cloning (Life Technologies).

Table 2.1. Primer sequences used for PCR amplification of the gene specific fragments used for the generation of the RNAi/ hairpin RNA binary constructs.

Primer	Sequence	Tm (oC)	Product size (bp)
KfMYB439RNAi F	CACCGTCATCATGTGTAACTGAACTCAGAC	61.6	442
KfMYB439RNAi R	CGTGTGCAGTTGCCATCTACTTCCTC	61.1	
KfCIBRNAiF	CACCAACAGGGAGTCCGCGCGCAG	65.9	251
KfCIBRNAiR	CTCTCCCCTACCACGACGGATAGC	62.5	
KfCDF2 RNAi F	CACCGCGAAAAGCTCAATATGGGAAAC	59.7	346
KfCDF2 RNAi R	CCTGAGGTTCTAAGTCGGGTGTTTTCG	61.3	

Red CACC represents the sequence added to forward primers for directional TOPO cloning into the pENTR/D TOPO vector.

2.1.5.3.2 Full-length open reading frame and HA-tagged full-length open reading frame primers

Primers were designed to contain the whole gene open reading frame (ORF; from the ATG start codon to the TAA or TAG stop codon) of the target gene (Table 2.2). The forward primer usually started, or included, the start codon ATG, and the reverse primer usually ended with or included the stop codon. For the HA-tagged constructs, in which the hemagglutinin HA-epitope peptide sequence was fused to the N- or C-terminus of the full length ORF, either the forward or reverse primer also contained the HA-tag sequence (Table 2.3). The CACC sequence was included in all forward primers to allow directional cloning into the pENTR/D TOPO-cloning vector.

Table 2.2. Primer sequences used for PCR amplification of the gene specific sequences cloned into the full-length over-expression binary constructs.

Primer	Sequence	Tm (°C)	Product size (bp)
KfMYB439FL F	CACCCGAACATCGAGAGCC	55.4	964
KfMYB439FL R	GCCTCGAAGGTAGTGTCCATCAG	58.8	
KfCIBFL F	CACCATGTGATGAGGTTCAGGA	54.8	517
KfCIBFL R	GACCGAATTAACCGAACCAA	50.7	
KfCDF2FL F	CACCGACTCTCAAGATGCTGCAGCTCAAAG	64.4	1404
KfCDF2FL R	CTAATCCCTGGTTCGTTCTATCCACG	59.5	

Red CACC represents the sequence added to forward primers for pENTR/D TOPO cloning.

Table 2.3. Primer sequences used for PCR amplification of the gene specific HA-tagged full-length over-expression binary constructs.

Gene	Primer Name	Sequence	T_m (°C)	Product size (bp)
CYCLING DOF FACTOR2	CDF2_CHiP_F_2	CACCATGCTGCAGCTCAAAGATCCGGCGATCAAGCTATTCGGC	69.7	1441
	CDF2_HATag_R_3	**TCA**AGCGTAATCTGGAACATCGTATGGGTAGCTCACTTCATGAAAATTCAAGGACCTTGATAAGGCTGCGGG	68.4	
	CDF2_ChIP_R_2	**TCA**GCTCACTTCATGAAAATTCAAGGACCTTGATAAGGCTGCGGG	69.1	1441
	CDF2_HATag_F_2	CACCATGTACCCATACGATGTTCCAGATTACGCTCTGCAGCTCAAAGATCCGGCGATCAAGCTATTCGGC	69	
CAM-INDUCED BZIP1	CIB1_CHiP_F_2	CACCATGTCAGGAGCGAACTATGTTGTCCCTGACAGCGCG	69	544
	CIB1_HATag_R_2	**CTA**AGCGTAATCTGGAACATCGTATGGGTACCACGGATAGCATAGGAGGTGGGGCTCTG	67.2	
	CIB1_CHiP_R	**CTA**CCACGGATAGCATAGGAGGTGGG	62.7	544
	CIB1_HATag_F	CACCATGTACCCATACGATGTTCCAGATTACGCTTCAGGAGCGAACTATGTTGTCCCTGAC	61.3	
MYB HOMEODOMAIN SUPERFAMILY PROTEIN	MYB439_CHiP_F	CACCATGGTTTCAGTGCGTACGAACAGCAGC	61.3	991
	MYB439_HATag_R_2	**TTA**AGCGTAATCTGGAACATCGTATGGGTATGTAGTCGAGATGGCACTTTGGTGG	59.3	
	MYB439_CHiP_R	**TTA**TGTAGTCGAGATGGCACTTTGGTGG	59.9	991
	MYB439_HATag_F	CACCATGTACCCATACGATGTTCCAGATTACGCTGTTTCAGTGCGTACGAACAGCAGC	59.1	

In the sequence the red CACC was added for directional cloning into the pENTR/D TOPO vector, the green ATG represents the gene's start codon, the bold black sequences represent the gene's stop codon, and the blue sequence (if present) represents the HA-Tag. T_m is calculated for the gene specific part of each primer, excluding the HA-Tag when present, The annealing temperature used in the PCR was between 2 - 5˚C lower than the highest T_m of each primer pair.

2.1.5.3.3 PCR fragment production

Using the specially designed primers for the PCR amplification of the gene specific fragments for the RNAi and full length over-expression constructs, a half reaction (25ul) KOD Hot start (Merck Millipore, Germany) PCR was carried out to produce blunt end PCR products suitable for pENTR/D TOPO cloning.

KOD Hot Start PCR Conditions:

Step	Temperature (˚C)	Time (s)
1 – Denaturing	94	120
2 – Denaturing	94	15
3 - Annealing	Lowest primer's T_m minus 2 to 5˚C	30
4 - Elongation	72	20s/1kb
Repeat Steps 2 – 4 for a further 29 cycles.		

10 µl of PCR product was separated and visualised using a 1 % (w/v) agarose gel containing 1xTAE, pH 8.5 (40 mM Tris acetate and 1 mM EDTA in reverse osmosis distilled water) and 0.1 µg/ml of ethidium bromide. If a single band was visualised then the QIAquick PCR Purification kit (Qiagen), was used to clean-up the DNA from the remaining 15 µl of PCR product. If more than one band was detected on the agarose gel, then the DNA band of the correct size was chosen and excised from the gel with a sterile scalpel, prior to clean-up using a QIAquick Gel Extraction Kit (Qiagen).

2.1.5.4 pENTR/D directional TOPO cloning

Cleaned-up PCR products, either from the QIAquick PCR purification kit or the QIAquick Gel Extraction kit, were measured using the Nanodrop spectrophotometer (Nanodrop 1000 Spectrophotometer) to determine the concentration of DNA

recovered for each PCR product. For blunt-end TOPO cloning of PCR products into the pENTR/D TOPO vector, 1 - 5 ng of a 1 kb PCR product, or 5 - 10 ng of a 2 kb PCR product was used in the TOPO cloning reaction, which was performed according to the manufacturer's instructions (pENTR/D-TOPO cloning kit, Life Technologies).

TOPO cloning reactions were performed in half volumes but all other procedures followed the manufacturer's instructions. The volume of the cleaned-up PCR product and the sterile molecular biology grade water was adjusted to achieve the correct concentration of PCR product in the TOPO cloning reaction.

Reagent	Volume (ul)
PCR Product	1*
Salt Solution	0.5
Sterile Water	1*
TOPO vector	0.5
Final Volume	3

This reaction was mixed and left at room temperature for 5 mins. 2 µl of the TOPO cloning reaction was added to a vial of TOP10 One Shot Chemically Competent *E.coli*, mixed gently, and left on ice for 30 mins. The cells were then subjected to a heat shock at 42˚C for 30 s, and then placed back on ice. 250 µl of Super Optimal broth medium with Catabolite repression (S.O.C; 2% tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM MgSO4, and 20 mM glucose) Medium was added, and the tubes were then left to shake at 195 rpm/37˚C for 1 h to allow recovery of the transformed cells. Various volumes were then plated onto LBA plates containing 50 µg/ml kanamycin for selection of E. coli cells harbouring the

pENTR/D plasmid. Plates were incubated in a walk-in growth room at 37˚C overnight to allow for visible single colonies to form.

2.1.5.5 Transformant analysis

After the overnight incubation, single isolated colonies were selected randomly and used for colony PCR using REDTaq (Sigma-Aldrich) to determine whether or not the desired insert was present in each colony. If a band of the correct size was amplified from the colonies, then a 10 ml overnight liquid culture was set up for each individual colony using LB containing 50 μg/ml kanamycin. Cultures were incubated in a shaking incubator overnight at 37˚C, 195 rpm. Plasmid DNA was isolated from the overnight liquid cultures using the QIAprep Spin Miniprep Kit, and was then submitted for commercial Sanger Dideoxy DNA sequencing using the GATC Biotech LIGHTrun service. Nucleotide sequences from each pENTR/D clone were aligned with the corresponding sequence from the assembled draft genome of either *K. fedtschenkoi* or *K. laxiflora*, using the bioinformatics software Geneious 4.5.5. Once a 100% match for each nucleotide sequence was found, that clone was taken forward for Gateway recombination of the insert sequence into the desired binary destination vector.

2.1.5.6 Gateway LR cloning

For the different binary constructs, different *Agrobacterium*-compatible, Gateway recombination cloning compatible binary vectors were used. For the RNAi hairpin-RNA binary constructs, the vector pK7GWIWG2(II) was used, whereas for constitutive over-expression of full-length ORFs under the control of the CaMV 35S promoter, vector pK2GW7,0 was used (Karimi *et al*., 2002). Following the Invitrogen/ Life Technologies method for LR Clonase II enzyme mix, the entry clone

and destination vector were mixed with the LR Clonase II enzyme mix, and incubated at 25˚C for 1 h. 10% (v/v) Proteinase K solution was added and the mixture was incubated for 10 mins at 37˚C terminate the reaction. Approximately 50 ng of the terminated destination vector/ entry vector LR Clonase II reaction was then added to Solo Pack Gold Competent *E.coli,* (Single-tube transformations (25ul); Stratagene) and the mix of cells and vector was subjected to a heat shock: 30 mins on ice, 30 s at 42˚C, followed by addition of 1 ml of SOC medium and incubation for 1 h at 37˚C with shaking. A range of volumes of the culture was then plated out onto antibiotic selection plates (LBA, with 300 µg ml^{-1} Streptomycin plus 100 µg ml^{-1} Spectinomycin), and incubated at 37˚C overnight. Transformant analysis was then repeated using colony PCR as described in section 2.1.5.5 above.

2.1.5.7 Producing competent *Agrobacterium tumefaciens* (Strain GV3101)

A small inoculum of ice/ culture scraped from a glycerol stock of *Agrobacterium tumefaciens* strain GV3101 stored at -80˚C was streaked out on LBA plates (25 µg ml^{-1} Gentomycin, 10 µg ml^{-1} Rifampicin) and grown overnight at 28˚C. Ten ml of selective LB liquid medium containing 25 µg ml^{-1} Gentomycin, 10 µg ml^{-1} Rifampicin was inoculated with a single *Agrobacterium tumefaciens* GV3101 colony from the overnight plate and grown overnight with shaking at 28˚C. The next day this starter culture was used to inoculate a 200ml of selective LB with the same antibiotics at the same concentrations. This was incubated at 28˚C with shaking until the absorbance of the culture at 600 nm (OD_{600}) reached between 0.5 and 1. The culture was then chilled on ice and centrifuged at 10,000 g for 10 mins at 4˚C to pellet the *Agrobacterium* cells. The supernatant was poured off into a waste beaker that was disposed of by autoclaving, and the pellet was resuspended gently in 1ml of ice-cold 20 mM calcium chloride, before re-centrifugation for 10 mins at 10,000 g, 4˚C. The

supernatant was then discarded by autoclaving and the cells were resuspended in ice-cold 20 mM calcium chloride (1 ml for every 250 ml of LB used for the large culture grown on day 2). 100 µl aliquots of the resuspended cells were transferred on ice into screw-cap 1.5 ml Eppendorf tubes, and the cells were snap frozen in liquid nitrogen and stored at -80˚C in a freezer until required for transformations.

2.1.5.8 *Agrobacterium tumefaciens* (Strain GV3101) transformation

1 µg of the desired binary construct plasmid DNA was added to 100 µl chemically competent *Agrobacterium tumefaciens* GV3101, and the mixture was left on ice for 15 mins. Cells were then snap frozen in liquid nitrogen, before being heated to 37˚C for 5 mins. 1 ml LB was then added for every 100 µl of *Agrobacterium* plus vector mix, and the culture was incubated for 2 h at 28˚C with shaking (250 rpm). The tubes were then centrifuged for 30 s at 21,000 g, and pellets were resuspended in 100 µl LB by gentle mixing. The 100 µl of resuspended cells was then plated out onto selective LBA plates (10 µg ml^{-1} Rifampicin, 25 µg ml^{-1} Gentamycin, 300 µg ml^{-1} Streptomycin, 100 µg ml^{-1} Spectinomycin). *Agrobacterium tumefaciens* colonies took 2 - 3 days to grow at 28˚C. Transformant analysis was then repeated as described in section 2.1.5.5 above, except that sequencing was not carried out as long as the band(s) obtained from the colony PCR reactions were of the correct size for the desired insert.

2.1.5.9 Stable transformation of *K. fedtschenkoi* and/ or *K. laxiflora* using *Agrobacterium tumefaciens* GV3101

Prior to *Kalanchoë* transformation with the confirmed binary constructs, overnight liquid cultures (30ml and 10ml in LB plus antibiotics) were grown at 28˚C with shaking (250 rpm) were grown in MS30 (Murashige and Skoog medium containing

Gamborg B5 vitamins 3% (w/v) sucrose,) until OD_{600} reached 0.1. 100 μM acetosyringone was added to induce the *Agrobacterium* virulence (*vir)* genes, and this mixture was left for 2 h at room temperature to allow virulence to develop.

Sterile leaf margin plantlets (for *K. fedtschenkoi*) or sterile seedlings (for *K. laxiflora*) that were ≥3mm (produced as described in Materials and Methods 2.1.5.1), were taken and cut into small 1 mm x 1 mm squares in a sterile flow hood using two sharp scalpels. Care was taken slice cleanly and not crush. The acetosyringone-induced *Agrobacterium* GV3101 culture harbouring the desired recombined binary vector carrying the target gene of interest was then poured into a sterile petri dish with the sliced plantlet/ seedling explants, and left to incubate for 1 h, with gentle agitation of the explants at regular intervals throughout the hour. After 1 h, the explants were removed from the *Agrobacterium* suspension and blotted dry on sterile filter paper. The explants were next placed onto petri dishes containing solid Callus Induction Medium (CIM: MS30 plus Gamborgs B5 plus 1 mg L^{-1} Thidiazuron (TDZ), 0.2 mg L^{-1} Indole-3-acetic acid (IAA) and 100 μM acetosyringone, solidified with 8% (w/v) Phytoagar). Plates were incubated for 48 h under 16:8 light :dark; 100 μmoles m^{-2} s^{-1}: 0 μmoles m^{-2} s^{-1}, 20˚C.

2.1.5.10 Tissue culture

After 48 h on the original CIM plates, the explants were transferred to fresh CIM plates (as described in Materials and Methods 2.1.5.9, no acetosyringone), with added antibiotics: 100 mg L^{-1} kanamycin (to select for plant cells transformed with the T-DNA from the binary vector) and 200 mg L^{-1} carbenicillin (to kill the *Agrobacterium* and prevent *Agrobacterium* over-growth). Once healthy green callus had formed with evidence of initial formation of vestigial shoot primordia, then the calli were moved

onto Shoot Induction Medium (SIM; MS30 with Gamborgs B5 vitamins plus 1 mg L^{-1} benzylaminopurine (BAP), 0.2 mg L^{-1} IAA, 100 mg L^{-1} kanamycin and 200 mg L^{-1} carbenicillin, solidified with 8% (w/v) Phytoagar). Once shoots had developed enough to allow the removal of a section of stem bearing leaves from the callus, then the base of the stem was cut with a sterile scalpel, and the excised shoot was placed into Root Induction Medium (RIM: MS30 plus Gamborgs B5 vitamins with 50 mg L^{-1} Kanamycin and 200 mg L^{-1} carbenicillin solidified with 8% (w/v) Phytoagar). Once roots had developed growing extensively into the kanamycin-containing selective medium, and if the leaves of the young plant remained completely green and healthy with its roots spreading through the 50 mg L^{-1} kanamycin RIM, then plants were removed from the agar plates, the excess growth medium was washed carefully from the roots, and the plantlets were placed onto soil. Plants were then left to grow under greenhouse conditions (16: 8 light: dark cycles, minimum temperature 20˚C) until they are large enough to be sampled for screening for the presence of the desired transgene.

Throughout tissue culture stages the growth conditions were kept the same: 16:8 light: dark, 100 µmoles m^{-2} s^{-1} : 0 µmoles/m^{-2}/s^{-1} , 20˚C, and were explants from the tissue culture plates were sub-cultured every two weeks onto freshly made plates (either CIM, SIM or RIM depending on the development of the explants).

2.2 RNA Isolation

2.2.1 RNA extraction kits

Leaf, stem, shoot apical meristem, bud or flower samples were ground in liquid nitrogen using a mortar and pestle. Total RNA was isolated using either the Qiagen RNeasy mini-kit (Qiagen), or the Spectrum plant total RNA kit (Sigma) according to the manufacturer's instructions. In addition to the manufacturer's instructions, 30 ul ml^{-1} of 50 mM PEG-20000 was added to the initial extraction buffer to remove phenolics compounds that interfere with efficient RNA isolation (Gehrig *et al*., 2012).

2.2.2 CTAB extraction method

Leaf samples were ground in liquid nitrogen using a mortar and pestle. RNA was then extracted using a modified version of Chang *et al* (1993). 10 ml extraction buffer containing 2 % β-mercaptoethanol was warmed to 65˚C and 2 – 3 g of each ground leaf sample was added, vortexed and incubated at 65˚C for 20 mins. Samples were then placed on ice and 10 ml chloroform:isoamylalcohol (24:1) was added, mixed well by inverting the tube several times, and then the tubes were centrifuged at 3600 x g for 10 minutes. The upper aqueous layer was carefully removed to a fresh tube. The chloroform:isoamylalcohol extraction was then repeated (two extractions in total). 0.25 volumes of ice-cold 10 M lithium chloride was then added to the recovered aqueous layer and mixed well immediately. This was placed at 4˚C overnight to precipitate the RNA. The RNA precipitate was pelleted by centrifuging the tubes for 30 minutes at 3600 x g. The pellet was kept, dissolved in 500 µl TE (10mM Tris and 1mM EDTA, pH8). A further clean up was performed by adding 500 µl chloroform, shaking well and spinning for 5 mins in a benchtop refrigerated microfuge at full speed, 4˚C. The upper aqueous layer was transferred to a sterile, RNase free 1.5 ml

Eppendorf tube. 50 µl of 3 M sodium acetate (pH 5.2) and 1650 µl of 100 % ice cold ethanol were added, mixed well and RNA was allowed to precipitate in a -80°C freezer for 1 h. The samples were then centrifuged at 10200 x g for 20 minutes at 4˚C, and the pellets were kept and dried using a Speed Vac Plus (Flowgen, GYROVAP, DNA concentrator) connected to a vacuum pump (Vacuubrand Gmbh and Co., Wertheim, Germany) and refrigerated condensation trap (Savant) that had been switched on at least 15 min before use. The dried pellets were then resuspended in 100 µl DEPC treated distilled water. All total RNA samples were stored at -80˚C.

2.2.3 Checking total RNA quantity and quality

The purity and concentration of the total RNA was determined using a Nanodrop spectrophotometer (Nanodrop 1000 Spectrophotometer) by measuring absorbance at 260 and 280nm. The software provided with the Nanodrop calculated the concentration of the RNA in each sample (µg ml^{-1}). Integrity of the total RNA was checked using a formaldehyde-MOPS agarose gels. The gels were 1.3% (w/v) agarose dissolved in 10x MOPS pH8, 10 % (v/v) formaldehyde and sterile water 80 % (v/v). Running buffer was 1x MOPS pH7. Sample buffer contained 23% (v/v) 10x MOPS pH8, 6.2% (v/v) formaldehyde, 70% (v/v) formamide and 0.1µg/ml of ethidium bromide. Total RNA samples (5 µg) were mixed well with loading buffer and heated to 65˚C for 2.5 minutes and immediately snap cooled on ice. The gel was subjected to electrophoresis for 2h at 100 V in a horizontal agarose gel apparatus. RNA samples were all diluted to 1 µg $µl^{-1}$ based on the concentration calculated from the Nanodrop A260 nm reading, before setting up reverse transcription (RT) reactions.

2.3 cDNA synthesis by Reverse Transcriptase

RT was carried out using the Qiagen QuantiTect Reverse Transcriptase kit according to the manufacturer's instructions. 1 μg of each total RNA was reverse transcribed in each reaction. The initial step removed contaminating genomic DNA from the intact RNA samples through incubation with Qiagen's proprietary gDNA Wipeout buffer at 42˚C for 2 mins. The second step achieved the reverse-transcription of the total RNA into cDNA using random priming and was carried out at 42˚C for 15 mins. Next, the reverse transcriptase was deactivated by heating the mixture to 95˚C for 3 mins. The resulting cDNA was diluted 1:5 using sterile RNase-free water.

2.4 PCR

2.4.1 Primer design

Oligonucleotide DNA primers were designed using the computer program Geneious Pro 4.5.5 based on cDNA contigs identified using BLAST searches against the *K. fedtschenkoi* draft genome assembly and/ or *de novo* transcriptome assembly generated from 454 sequencing data, or using the draft genome and transcriptome assemblies for *K. laxiflora* for the HA-tagged binary constructs. Once designed, primers were synthesised commercially by Integrated DNA Technologies. Primers were diluted to 100 μM. Average product sizes for semi-quantitative RT-PCR ranged from 300-1500bp. For Q-RT-PCR analysis, primers were re-designed to produce an average product size of 100 bp.

2.4.2 RT-PCR

1 µL of each diluted RT cDNA sample was used in each PCR reaction (10 µL) containing 1x Sigma REDTaq ReadyMix PCR reaction mix with $MgCl_2$ (product number R2523) and 1 µL of each forward and reverse primer (final concentration 1 µM). PCR reactions were conducted following standard PCR protocol described by Boxall *et al.* (2005). Optimal cycles and annealing temperatures varied for different primers (Table 1). The housekeeping gene Polyubiquitin10 (*UBQ*; KfUBQ10, GenBank accession KM114222) was used as a loading control and all values for band intensities were normalized to the *KfUBQ10* gene.

Step	Temperature (°C)	Time - Seconds
1 – Denaturing	95	120
2 – Denaturing	95	15
3 - Annealing	Lowest primer's T_m	30
4 - Elongation	72	20s/1kb
Repeat Steps 2 – 4 another x cycles. Varies depending on transcript being amplified.		

Table 2.4. Primers used for RT-PCR.

Gene/ Primer Name	Forward Primer (5^1-3^1)	Reverse Primer (5^1-3^1)	T_A for PCR reaction (°C)	PCR Cycles used
CDF2 RNAi	GGCCTCGCCGGGGATCATTC	TCTTACGCCTGCCGGACCA	57	26
CDF2 FL	CACCGACTCTCAAGATGCTGCAGCTCAAAG	CTAATCCCTGGTTCGTTCTATCCACG	57	26
CIB	CACCAACAGGGAGTCCGCGCGCAG	CTCTCCCCTACCACGACGGATAGC	64	27
MYB 439	GGCGGGGCTGGAGGTTGGAT	CCGGGAGAGCTGGACCGACA	57	26
UBQ10	ACACCATCGACAACGTCAAA	CACCTCAGGGGTGATGGTCT	51	24
TOC1	CCTTTTGCGCAATATCCT	ACGCGTCCCTAAAAGTTCAC	52	28
CCA1	ATTTTCATCATCGGCTGGTC	GTTGCAAGGGGTAGCTGGTA	55	40
PEPC	GGCCTACACCCTGAAACGTA	CAGTAAGCTCCCTCCTGCTG	60	22
PPCK	CACCGCGCAGGACATATTTGAGG	GGCCGAATTCCCTCACTCTGCTTC	61	26
NAD-ME1 alpha	CACCATCAGTCAACAACTAGTCCATCTTTC	CCATTATAAGCCACTAACAGGAGGAGG	63	24
NAD-ME1 beta	CCCTTCCTCAGCCTGCTGTGTAATTG	CCGAATACCACAAGACAGAATGTAGTGATC	63	24
PPDK	TTGGGACGAATGATCTCACA	CCATCTCCCACATCCACTTT	19	55
PPDK-RP	AGCCGAGCTCCCCTTTAAGTAAGCTG	AAGGTCGCCTGCTTAACAATCTCCAT	63	27
GPT2	CACCATTGACATTCAGCATCGGTAAC	GGGATGGAACAATTAGTTGCAGCAG	54	24
MEX1	CACCGCTGTCCTACTTTGTTGTCATCTCATC	GAGCGGATAATCCTTTGATGTTACTCGG	56	24
GWD	CACCGGTGCTGGCCTCTACGACAG	CGATCTCCCTGGCAGCACTGGTG	63	22

2.4.3 Visualisation of semi-quantitative RT-PCR products

1 % (w/v) agarose gels containing 1xTAE, pH 8.5 (40 mM Tris-acetate and 1 mM EDTA in RO water) and 0.1 μg ml^{-1} of ethidium bromide were used to separate and visualise PCR products. Running buffer contained 1x TAE and 0.1μg/ml of ethidium bromide. All gels were subjected to electrophoresis for 35 minutes at 100 V. Gels were then imaged using a UV light box integrated into a Geneflash imaging system (Syngene Bio-imaging).

2.4.4 Semi-quantitative RT-PCR image analysis

To analyse band intensities, Metamorph (Meta Imaging Series 6.1) was used. Each band was measured using "regional properties" and a box was drawn around each band, which is automatically numbered. 4 background readings were also taken for each gel to account for gel differences. Polyubiquitin10 (*KfUBQ10*, Genbank accession KM114222) intensities were measured as a reference gene that maintains stable transcript abundance under the conditions being studied. This allowed correction of the gene specific band intensities for minor variations in cDNA levels between samples. A normalisation factor was calculated for each cDNA sample based on the *KfUBQ10* band intensity. Gene specific band intensities were then measured and the normalisation factor was used to calculate the gene expression levels for transgenic and wild type plants. Averages were taken for the biological replicates, and graphs were made using Excel using averages values, with error bars representing the standard errors calculated from the three biological replicates.

2.4.5 RT-qPCR

1 µl of each diluted RT cDNA sample was used in each PCR well. Reactions had a final volume of 10 µl including 1x NO ROX SYBR Green mastermix (Bioline; BIO-98005) and 0.4 µl of each primer (final concentration 0.4 µM).

Table 2.5. Primers used for RT-qPCR. All primer sets were run at T_A 60˚C for 40 cycles.

Gene	Primer Name	Forward Primer	Reverse Primer
CYCLING DOF FACTOR2	CDF2	ATA CGA ATC GTG GAT AGA ACG AAC CAG G	CCT ATC TTC CTG AGG TTC TAA GTC GGG
APETALLA1	AP1	CAC CTC TAC ATA CTT CTA CAC TAC AGC C	GGG CTA GAC ACA ATC CAT ATA GAG ATC C
CONSTANS	CO2581	CGG CTC AGA GGC CGT GTC G	GCC CGA ACC CGG AAA AAA GTA AGA TGG
	C04615	TTC CAC GGC TAA GAT GCT CAG TCG	CCA GAA GCC CGA AAA AAA GTT GAA GGG
FD	FD	AAA TGC TTC CTG CCT GTG CTA TAC G	GGC TCT ATA TCT CCT GCT CAC CC
FLOWERING LOCUS T	FT12903	GGA GGA GAA TAA TAT CCG TGG CCA CG	CGT TAA GAG TCA CTT GGT GAA ATG TTT GGC
	FT18180	GCG CGC TGC GCA TCG ACG	GAA GGC GTC GAT CAC GTC TCC C
FRUITFUL	FUL	TTG GTA CTC ATA AGT CCT ACC TGC G	CCC GAT ACG AAC TAC TCC CAA TGC
GIGANTEA	GI	GCT TGA CAT GGG AAG CTC ACA ACC	CGT CAT ATC CTC ATA CTC ACA CAA ACC C
SEPALLATA1	SEP1	TTG GGA GTG GCA AGA TTT CAC TGG C	CCC TAG ACC AAA CAT GGG TGG TCC
SUPRESSER OF OVEREXPRESSION OF CO1	SOC1	GTT GAA GTG ATC AAC TAG CAG CAG AAG C	GGG CTG TGG GTT TGG AAG ATG ATG ATG
ZEITLUPE	ZTL	TCC CGG GGC AGG GAC AAG C	ACT TGT GCC CCA ATT CAT TCT CAA AAT CC
CONTIG9471 (reference gene)	9471	TGT TGT TCC TGC CAC	TAG AGG GTG AAG GTC

Before carrying out RT-qPCR assays on experimental time course samples, primers were tested for efficiency. All primer sets used produced one peak in dissociation curves, had R^2 values of between 1.0 and 0.99 and efficiencies between 99.8% and 116% (Appendix: S2.1-S2.12)

Q-RT-PCR reactions were conducted using the following standard PCR protocol.

Step	Temperature (°C)	Time - Second
1 – Denaturing	95	120
2 – Denaturing	95	5
3 - Annealing	60	10
4 - Elongation	72	10
Repeat Steps 2 – 4 for 40 cycles.		
5 – Dissociation Curve:		
Denature	95	30
Anneal	60	30
Elongation	72	30

2.5 Leaf disc pH screens

All plants were entrained to 16:8 light: dark cycles (100 µmol m^{-2} s^{-1}: 0 µmol m^{-2} s^{-1}, 25˚C: 15˚C, 60%: 80% humidity) in a plant growth chamber (Snijders Scientific, Microclima MC1000). Leaf pH assays were conducted on each plant by collecting three replicate 2cm^2 leaf disc samples from leaf pair 6 of each plant both 1 h before dawn and 1 h before dusk. The discs were immediately immersed in a 0.025% (w/v) chlorophenol red solution (Sigma-Aldrich). Chlorophenol red is a pH indicator dye, which is yellow at pH4 and red at pH6. After incubation of the leaf discs in the chlorophenol red solution for 4 h, the absorbance of the chlorophenol red was measured at 574 nm using a SpectraMAX340 plate reader (Molecular Devices) and SOFTmax PRO software to enable comparisons between the wild type and transgenic plants.

2.6 Iodine staining to estimate leaf starch content in leaf discs

The same plants used for the acidity staining described in 2.5 were also sampled simultaneously for iodine staining to test for leaf starch content. Three 2 cm^2 leaf disc replicates from leaf pair 6 were sampled at 1 h before dusk and 1 h before dawn and cleared of chlorophyll and other leaf pigments in 100 % ethanol for approximately 5 days; by which time they were white. The discs were then immersed in a starch staining solution consisting of 5.7 mM iodine, 43.4 mM potassium iodide and 0.2 N hydrochloric acid (HCl). After 20 mins staining, the solution was removed and the discs were washed with 80% (v/v) ethanol and imaged. Starch containing leaves turned purple/black, whereas leaf discs with low levels of starch developed a more yellow/ brown colouration.

2.7 Gas exchange measurements

2.7.1 Multi-Channel IRGA

Rates of net CO_2 assimilation in *K. fedtschenkoi* leaf pair 6 were measured continuously for two weeks (7 days 12:12 light dark, 400 µmol m^{-2} s^{-1}: 0 µmol m^{-2} s^{-1}, 25˚C: 15˚C, 60 %: 70 % humidity, 7 days constant light, temperature and humidity: 100 µmol m^{-2} s^{-1}, 15˚C, 70 % humidity) using a bespoke 6-cuvette Infra-Red Gas Analysis system (IRGA). This multi-channel gas exchange system was constructed by Dr. Keith Parkinson of PP Systems Ltd (Hitchin, Hertfordshire) for the Hartwell lab (University of Liverpool), and is described in full detail in Dall'omo (2011) and Dever *et al.* (2015).

Before leaves were removed from the plant and placed with their petiole in distilled water in the gas exchange cuvettes, the plants were first entrained to 12:12 light: dark cycles (400 µmol m^{-2} s^{-1}: 0 µmol m^{-2} s^{-1}, 25˚C: 15˚C, 60 %: 70 % humidity) in a plant

growth chamber (Snijders Scientific, Microclima MC-1000). Leaf pair 6 were excised at their petiole from their parent plant with a sterile scalpel, and placed immediately in 150 ml beakers of distilled water with their leaf petioles submerged in order to avoid cavitation. Each beaker of water was covered in several layers of parafilm and small incisions through the parafilm were made with a scalpel in order to create a small hole for the leaf petiole to reach through into the distilled water below. The parafilm minimised the evaporation of the water into the beakers into the cuvette atmosphere, ensuring that the major changes to the atmosphere in the gas exchange cuvette were due to changes in the physiology of the leaves, such as stomatal opening leading to CO_2 fixation and transpirational water loss from the leaf surface.

To account for different leaf areas, at the end of each gas exchange experiment, leaves were scanned on a standard flat-bed scanner and Image J freeware image analysis software was used to analyse the images pixels of each leaf and produce a leaf area. Each leaf area value was then entered into a "recalculate photosynthesis values" feature within CIRAS SC-DC Control Software (V1.03 [Build 07]). This feature recalculated the actual photosynthetic rate in "μmoles CO_2 fixed (or respired) per m^2 leaf area per second" for each cuvette at each measured time point over the duration of the time course experiment. The area under the curve was also calculated for some of the gas exchange datasets in order to determine the total amount of CO_2 fixed and/or released per 12 h light or dark period during both LD and LL conditions. This was achieved using an Excel macro script designed by Ben Wareham which allowed the calculation of the peak area according to Simpson's rule.

2.7.2 LICOR Single-leaf IRGA: LI-6400XT

Rates of net CO_2 assimilation, stomatal conductance and intercellular CO_2 concentration in *K. fedtschenkoi* LP7 were measured using a LICOR LI-6400XT portable photosynthesis system, which included a fluorescence head (Product number:6400-40) (LI-COR Biosciences, Nebraska). A custom 30 h programme was designed by LI-COR to allow the measurement of K. fedtschenkoi or K. laxiflora leaf gas exchange over a 30 h period. This programme used the external light sensor on the IRGA head to track the light conditions in the growth chamber. When light intensity in the chamber was higher than 200 μmol m^{-2} s^{-1}, then the LED lights in the IRGA head were turned on and the light intensity inside the head was matched to track the chamber light intensity. Below 200 μmol m^{-2} s^{-1}, the LED lights in the IRGA gas exchange cuvette were off. External conditions in the Snijders Microclima MC-1000 growth cabinet in which the LICOR measurements were undertaken were set to 12:12 light: dark cycles (400 μmol m^{-2} s^{-1}: 0 μmol m^{-2} s^{-1}, 25˚C: 15˚C, 60 %: 70 % humidity.

In the clamp-on leaf gas exchange cuvette that is used for gas exchange measurements with the LICOR LI-6400XT, the air flow-rate was set to 300 μmol s^{-1} due to the fluorescence head being attached. The CO_2 concentration inside the cuvette was set to 400 μmol mol^{-1} (equivalent to 400 ppm), with the CO_2 mixer turned on. Each attached leaf (still actively growing as part of its mother plant) was then sealed into the IRGA gas exchange cuvette by clamping it gently onto the leaf with sufficient pressure to farm an air-tight seal, but without damaging the leaf. The cuvette was supported so that it didn't place strain on the leaf or its petiole. The leaf area that fitted into the gas exchange cuvette was a fixed 2 cm^2 circle, and this area was programmed into the machine.

2.8 Leaf metabolite assays

Leaf pair 6 from mature plants were sampled after plants were entrained to 12:12 LD conditions for 7 days as described in Methods 2.1. Leaves were collected 1 h before dawn and dusk, and immediately placed into liquid nitrogen and then stored at -80˚C. Samples were ground to a fine powder in liquid nitrogen using a pestle and mortar. A K-factor was produced for each of these assay buffers to enable path length correction. The spectrophotometer used for all of these assays was a MultiScan GO 1510 (Thermo Scientific).

2.8.1 Malate assays

Approximately 0.5 g of ground, frozen leaf tissue was added to 4 ml of 80 % methanol (v/v) and heated in a water bath at 80˚C for one hour with intermittent agitation. Samples were then centrifuged at 3636 g for 10 mins and the supernatant removed and stored on ice. This procedure was then repeated using 2 ml and 1 ml 80 % methanol and supernatants were pooled together to generate the total soluble metabolites extract for each leaf sample.

The pooled metabolite extract for each leaf sample was dried down using a Speedyvaccum concentrator (Savant SPD1010, Thermo Scientific) at 60˚C. Once the methanol had completely evaporated (after approximately three hours) the dry residue was re-dissolved in 0.5 ml of 0.2 M Bicine buffer pH 7.8. This was then centrifuged briefly in a microfuge to remove all insoluble material.

Malate was measured using an enzyme-linked spectrophotometric assay as described by Mollering (1974). The assay mix consisted of 76 mM 3-Amino-1-Propanol pH 10.0, 50 mM L-Glutamate pH10, 2 mM NAD, 1 U Malate Dehydrogenase, 20 U

Glutamate-oxaloacetate aminotransferase. A K-factor was calculated so that pathlength correction could be applied to the results. The assay was read at 340nm.

$$\mathbf{C = \frac{\Delta A}{\varepsilon} * V * v \ / g = mmol\ malate\ /\ g\ fresh\ weight}$$

ε = absorption coefficient (6.22), ΔA = absorbance change, V = total assay volume, v = volume of Bicine buffer the sample was dissolved in, g = weight of tissue used

2.8.2 Starch assays

The method used for starch assays followed the method described by Smith and Zeeman (2006). Starch was extracted from the insoluble material remaining after the centrifugation steps during the hot, 80 %-methanol extraction of soluble leaf metabolites (see 2.8.1). The pellets were heated in a water bath for about 5 mins until they were dry as a result of the evaporation of all residual 80 % methanol. The pellet was then re-dissolved in 5 ml of MilliQ Water, and vortexed vigorously until the pellet had been disrupted.. 0.5 ml of each starch solution was placed into three tubes, two to be digested by Amyloglucosidase (6 U per tube) and alpha-amylase (1 U per tube) enzymes, and one to be the control (MilliQ water was added instead). These tubes were then left overnight at 37˚C with constant shaking, so that all of the starch could be hydrolysed to glucose. The solution of glucose was then measured with an enzyme linked spectrophotometric assay, using Hexokinase and Glucose-6-phosphate dehydrogenase (0.2 untis of each). The reaction was measured at 340 nm (Smith and Zeeman, 2006). Path-length was again corrected for using the K-factor of 0.176.

$$C = \frac{\Delta A}{\varepsilon} * V * d \ / g * MW = \text{mol starch / g fresh weight}$$

ε = absorption coefficient (2.91), ΔA = absorbance change, V = total assay volume, d = dilution of sample in assay, g = weight of tissue used, MW = molecular weight of starch.

2.8.3 Sugars –glucose, fructose and sucrose- assays

Three sugars were measured using a coupled spectrophotometric assay that allowed the determination of the level of D-Glucose, D-Fructose and Sucrose in the soluble metabolite extracts generated in 2.8.1. Using a series of enzymatic reactions, listed below, all three sugars could be quantified using a spectrophotometric assay at 340nm at 25˚C. The assay mix consisted of 10 mM imidazole pH6.9, 5 mM magnesium chloride, 1 mM ATP, 2 mM NADP and 0.4 U G6PDH per 200 µl reaction. Standard reactions were 10 µl metabolite extract (in 0.2 M Bicine; 2.8.1) plus 190 µl of assay mix. This mix was measured at 340 nm for 10 mins in the microplate reader to generatre a zero-point absorbance reading. All free Glucose-6-Phosphate (G6P) during this 10 minute incubation period will be converted to Gluconate-6-Phosphate (Glu6P).

After 10 mins 0.3 U Hexokinase (HK) was added to 3 wells (and water added to the control well). This was mixed and incubated for 30 mins at 35˚C, so that all of the D-Glucose present in the metabolite extract can be converted to G6P, which can then be converted to Glu6P by G6PDH generating NADPH that is proportional to the amount of glucose in the original extract.

Once this reaction had finished (as determined by a stable A_{340nm} reading), 0.3 U of phosphoglucose isomerase (PGI) was added. This converted the fructose-6-phosphate

(F6P), generated from D-Fructose by HK in the previous step, to G6P, and then the pre-existing G6PDH in the reactions converted the G6P to Glu6P, generating further NADPH and thus further increasing the A_{340nm} reading by an amount that was proportional to the amount of fructose in the original extract.

Finally, to determine the level of sucrose in the extract, 0.8 U invertase was added when the previous reaction had run to completion (as judged by a stable A_{340nm} reading) after about 30 mins. This step converted sucrose to glucose and fructose, which were then acted on by the pre-existing HK and PGI in the assay mix to generate G6P. G6PDH again acted to convert G6P to Glu6P, thereby further increasing the A_{340nm} value by an amount that was proportional to the amount of sucrose in the original extract.

Absorbance increases at the end of each step of the reaction were calculated, which allowed the amounts of glucose, fructose and sucrose in the original extract to be determined using the following equation:

$$\mathbf{C = \frac{V * MW}{e * v * 1000} *\Delta A = g\ sugar/L}$$

V=final volume (ml), MW=Molecular weight, e=NADPH absorption coefficient at 340nm (6.3), v=sample volume (ml).

2.9 Leaf protein extractions and immunoblotting

Total protein was extracted from *K. fedtschenkoi* leaf tissues by grinding approximately 0.5 g of ground leaf tissue with a small quantity of acid washed sand and then adding a 1x SDS sample buffer (1 M Tris-HCL, pH6.8, 20 % v/v glycerol, 4 % v/v 2-mercaptoethanol and 3 % w/v SDS). This was all then transferred to a screw cap tube and boiled from 10 mins. Samples were then allowed to cool to room temperature, and then the pH was checked, and adjusted if necessary, to pH 6.8. Samples were then centrifuged for 5 mins at full speed in a benchtop microfuge. The supernatant was kept, and a sub-sample used to determine protein concentration, using Bradford Ultra Reagent (Expedion). The protein concentration was used to determine the amount of sample that should be loaded onto each lane of the SDS-PAGE gels according to the details for each protein that was to be detected using immunoblotting in Table 2.6 below.

Table 2.6. Proteins detected using antibodies for immunoblotting, with the amount of protein required, and dilutions of the antibodies.

Protein being detected	µg protein required for detection	Dilution of primary antibody	Dilution of secondary antibody
PEPC	5	1:5,000	1:5,000
Phospho-PEPC	20	1:1000	1:20,000
NAD-ME α and β	15	1:100	1:10,000
PPDK	10	1:10,000	1:10,000
PPDK-RP	20	1:3,000	1:5,000
CIB1	40	1:100	1:5000

Proteins were separated using one-dimensional 12 % SDS-PAGE gels. Gels were subjected to electrophoresis for approximately 1 h at 150 V, then either stained in Coomassie Blue, or blotted to a nitrocellulose membrane for 1 h at 100 V in a 4˚C

room. Ponceau S stain was then used to confirm protein transfer, and to check for even loading of protein in each lane of the gel. The blot was then blocked overnight in 5 % (w/v) dried milk powder. Primary antibodies used were as follows:

Table 2.7. Primary antibodies used for Immunoblotting, stating the plant which had the antibody generated to it and how the antibody was obtained.

Protein detected by anitbody	Antibody raised against	Supplied by	References
PEPC	*K. fedtschenkoi*	Hugh G. Nimmo, University of Glasgow	(Nimmo *et al.,* 1986)
Phosphorylated PEPC	*Hordeum vulgare* (Barley)	Cristina Echevarria, Universidad de Sevilla	(Gonzalez et al., 2002; Feria et al., 2008)
PPDK and Phosphorylated PPDK-RP	*Zea mays* (Maize)	Chris J. Chastain, Minnesota State University	(Chastain et al., 2000, 2002)
NAD-ME α and β	*A. thaliana*	Maria F. Drincovich, Universidad Nacional de Rosario	(Tronconi et al., 2008)
CIB1 (C terminal)	*K. fedschenkoi* peptide sequence – commercially synthesised.	GenScript USA Inc.	(Hartwell et al., unpublished)

Secondary Horse Radish Peroxidase (HRP)-linked rabbit IgG antibody (from donkey – GW Healthcare, RPN2108) was used. Dilution of the secondary antibody varied depending on the protein being detected (Table 2.6). Blots were then developed using the ECL system (GE Healthcare). They were incubated in the solution for 1 min. Excess solution was then removed carefully. Exposure times for X-ray film placed over the blot in the X-Ray film cassette varied between 30 s for PEPC to 10 mins for CIB. After exposure, the X-ray film was developed, then placed in fixative and then

finally washed. Film was scanned in order to generate the electronic versions of the blots included as results.

2.10 Enzyme assays

For all enzyme assays, leaf tissue was sampled and immediately frozen in liquid nitrogen. Samples were then ground in pestle and mortar with acid wash sand and the specific enzyme extraction buffer, with approximately 0.5 g tissue to 1.5 ml extraction buffer. The spectrophotometer used for all assays was a MultiSkan GO 1510 plate reader (Thermo Scientific), and all assays were performed in a 96-well plate format.

2.10.1 PEPC activity assay

The extraction buffer used contained; 50 mM Tris (pH8.0), 2 mM ethylenediaminetetraacetic acid (EDTA), 10 mM L-malate, 1 mM benzamidine hydrochloride (BZH), 1 mM dithiothreitol (DTT), 2 % (w/v) polyethylene glycol 20000 (PEG20000), 100 mg/g tissue sodium bicarbonate ($NaHCO_3$), and 200 mg/g tissue polyvinylpolypyrrolidone (PVPP). Assay buffer consisted of; 50 mM Tris (pH7.8), 5 mM magnesium chloride ($MgCl_2$), 2 mM phosphenolpyruvate (PEP), 0.2 M reduced nicotinamide adenine dinucleotide (NADH), 10 mM $NaHCO_3$ and 5 U malate dehydrogenase (MDH; pig heart).

The samples used in the PEPC assays were collected at 6 h into the 12 h dark period from mature plants entrained in 12:12 light: dark cycles in the Snijders Microclima MC-1000 growth cabinet. Frozen samples were ground rapidly in the extraction buffer with a small amount of added acid washed sand, and, once defrosted, they were filtered through Miracloth (Merck). A small amount (200 µl) of the sample collected was then removed for chlorophyll determination and stored on ice. The rest was then

centrifuged for 10 s at 21,000 g in a microfuge, and the supernatant was transferred to another tube.

From this extract, 10 µl was added to 190 µl of assay buffer, and the reaction was measured at 340 nm until complete (every 20seconds for 10minutes). The maximum rate between 3 points during the reaction was then calculated using the SkanIt RE 3.2 software for Multiskan GO.

Chlorophyll determination was then carried out by adding 80% (v/v) acetone to the sample kept without centrifuging in a 1:5 Sample:80% acetone ratio. The value was read at 652nm as detailed in Arnon (1948). A Bradfords assay was also carried out on the sample that was used in the assay. This enabled PEPC enzyme activity to be expressed in both µmol min^{-1} mg^{-1} chlorophyll and µmol min^{-1} mg^{-1} total protein.

2.10.2 Determination of the apparent K_i of PEPC for malate

The extraction buffer used for the measurement of the apparent K_i of PEPC for malate was the same as used for the PEPC assay (Materials and Methods 2.10.1); 50 mM Tris (pH8.0), 2 mM EDTA, 10 mM L-malate, 1 mM BZH, 1 mM DTT, 2 % (w/v) PEG20000, 100 mg/g tissue $NaHCO_3$, and 200 mg/g tissue PVPP. Samples were collected at 2 h into the 12 h light period and 6 h into the 12 h dark period mature plants entrained under the standard 12:12 light: dark conditions describer in 2.?. Frozen leaf samples were ground in extraction buffer and filtered through Miracloth (Merck) as in the PEPC assay, and were spun at 21,000 g for 10 s, before being stored on ice. These samples were then de-salted by running them through a desalt column using Sephradex-G25 medium (Sigma-Aldrich) that was left to swell overnight in RO water at 4^{o}C. The Sephadex-G25 columns (Soda-lime glass Pasteur pipettes with short capillary tip. Small amount of glass wool in tip and 2ml of swollen Sephradex-

G25 medium added on top) were equilibrated with 3 volumes of ice-chilled desalt buffer (50 mM Tris (pH 7.8), 5 % (v/v) glycerol, 1 mM DTT and 1 mM BZH) before samples were added. When washing the columns between reuse 0.2 M Sodium Hydroxide (NaOH) was used. Samples were collected in 0.5 ml fractions and the most concentrated sample (judged by colour intensity) was then used for the assay. The assay buffer used contained 50 mM Tris (pH7.8), 5 mM $MgCl_2$, 10 mM $NaHCO_3$, 1 U ml^{-1} MDH with MilliQ water added to bring the volume up to the final assay volume. Assays were performed in a 96-well microtitre plate using a final volume of 200 µl of assay buffer in each well. 6 µl of each sample and 1 µl 0.2mM PEP were added to start the reaction. The plate was incubated for 30 mins at 30˚C to allow the enzyme reaction to begin and reach a steady rate. After this time varying amounts of L-malate were added to different wells containing the same desalted extract to achieve final L-malate concentrations in each well between 0 mM to 10 mM. The maximum rate of the PEPC both before and after malate addition was calculated using the SkanIt RE 3.2 software, and the percentage inhibition for the different concentrations of malate was calculated in order to find the concentration of malate that caused 50% inhibition of the enzyme's maximum rate (K_i).

2.11 Experiments to study the response of transgenic *K. fedtschenkoi* lines to drought treatment

2.11.1 Collecting samples

Plants for the drought stress experiment were grown from developmentally synchronised leaf margin adventitious plantlets, and were 6-months-old when the experiment commenced. All plants were grown in greenhouse conditions (16:8 light: dark, approximately 250 μmoles m^{-2} s^{-1}: 0 μmoles m^{-2} s^{-1}, 23˚C: 18˚C), and the drought treatment and control were also conducted under the same greenhouse conditions. A total of 191 plants from wild type, and the best *KfCDF2*, *KfCIB1* and *KfMYB439* transgenic lines, both over-expressor and RNAi, were included. For each genotype, half of the plants were maintained well-watered (control), and half were subjected to progressive drought stress by completely withholding water. At least 6 clonal and developmentally synchronised plants were subjected to both conditions. Plants were left in their respective conditions for 31 days. A subset of each line (Table 2.8) in both drought-stress and well-watered conditions were moved to the Snijders Microclima MC-1000 growth cabinets under standard 16:8 light: dark conditions for the last 7 days of the experiment to enable leaf pair 6 (LP6) sample collection at 1 h before dawn and 1 h before dusk for malate and starch assays.

Table 2.8. Number of plants for each line subjected to each growth condition.

Lines	Number of plant in well-watered conditions	Number of plants in drought conditions
WT	8	8
MYB439 FL 37C	6	6
MYB439 FL 54B	9	9
MYB439 RNAi 3C	7	7
MYB439 RNAi 10B	9	9
CIB1 FL 5B	8	8
CIB1 FL 11B	11	11
CIB1 RNAi 12A	6	6
CIB1 RNAi 30A	7	7
CDF2 FL 13A	6	7
CDF2 FL 14C	9	9
CDF2 RNAi 19A	8	8
CDF2 RNAi 26B	9	9

Sampling of the plants maintained in the greenhouse throughout the experiment was carried out after 31 days, and photographs were taken of each line for comparison. Two 2.0 cm^2 leaf discs were collected from leaf pair 6 per plant and immediately frozen in liquid nitrogen. These discs were used to quantify chlorophylls A and B, plus anthocyanins and carotenoids. The remainder of the above ground material was weighed immediately to determine fresh or wet weight. The material was then bagged and dried in an oven at approximately 70˚C until the bag reached a constant weight. This allowed total dry weight of above ground material to be determined. Subtracting the dry weight from the wet weight allowed the determination of the percentage moisture content of each sample. In order to weigh the roots of each plant, all of the soil had to be removed. For well-watered plants root washes were carried out, dipping the soil into water until only the roots were left. These were dabbed dry to remove excess water and weighed. The drought-stressed roots had such dry soil that it was

easily removed by hand. Both sets of roots were then weighed for fresh/ wet weight, and then dried in an oven and reweighed to determine dry weight.

2.11.2 Moisture quantification

During sampling the moisture level of the soil in mV was measured using a soil moisture meter HH2 (Delta-T devices, Cambridge). An initial measurement was taken on the first day of the experiment, and a final measurement was taken after 31 days, enabling quantification of the moisture level in each plant's soil. The programme used to quantify soil moisture content was designed by Raj Whitlock, and allowed the calculation of the average volumetric water content of the soil (kg L^{-1}) for both the well-watered and the drought-stressed lines.

2.11.3 Malate and starch assays on well-watered and drought-stressed samples

Assays for malate and starch were carried out as described in Materials and Methods 2.8.1 and 2.8.2.

2.11.3.1 Total leaf anthocyanin content

Assays for leaf total anthocyanin content were carried out by extraction of 2 cm^2 of frozen, ground leaf tissue in acidified methanol (methanol: HCL: water, 90:1:1). This was centrifuged for 30 s at 21,000g in a benchtop microfuge, and the supernatant was transferred to a fresh tube. According to the method described by Sims & Gamon (2002), if the absorbance was over 1 when measured at 529nm or 650nm then the sample should be further diluted in acidified methanol. For these samples, an extra 12 ml of acidified methanol was added. Anthocyanin content was calculated using this equation:

$$\textbf{Anthocyanin} = A_{529} - (0.24 * A_{650})$$

From this, the concentration of total anthocyanins in each leaf disc could be calculated using the Molar extinction coefficient of anthocyanins, which is 30,000, to calculate the moles of anthocyanin/cm^2 leaf.

2.11.3.2 Assays for Chlorophyll A and B and Carotenoids

Assays for chlorophyll A and B and carotenoid pigments were carried out using a 2.0 cm^2 leaf disc sampled from leaf pair 6, which was ground in liquid nitrogen before adding 2 ml of ice-cold acetone/1M Tris buffer was added (80:20 (v:v) pH7.8). This extract was centrifuged for 30 s at 21,000 g in a benchtop microfuge, and the supernatant was transferred to a fresh tube. Again, according to the method described by Sims & Gamon (2002), if the absorbance was over 1 when measured at 537 nm, 647 nm or 663 nm, then the sample was further diluted in acetone/Tris buffer. For these samples, a further 15.5 ml of the acetone/ Tris buffer had to be added. Each sample was then measured at 470nm, 537 nm, 647 nm and 663 nm. Chlorophyll A and B and total carotenoid content were calculated using equations defined in Sims & Gamon (2002), providing final values for each pigment in μmol/ml:

Chlorophyll A (Chl_a)= $0.01373*A_{663} - 0.000897*A_{537} - 0.003046*A_{647}$.

Chlorophyll B (Chl_b) = $0.02405*A_{647} – 0.004305*A_{537} – 0.005507*A_{663}$.

Carotenoids = (A470 – (17.1* (Chl_a + Chl_b) – 9.479* Anthocyanin)) / 119.26

Dilutions were then taken into account.

2.11.4 Measurement of leaf succulence

Twenty-four individual clonal, developmentally synchronised plants of both the wild type and each *KfCDF2* line (FL 13A, 14C and RNAi 19A, 26B) were grown in 16:8 light: dark greenhouse conditions for 6 months. All plants were watered daily. A soil moisture reading was then taken at day 0, and three plants from each line that were representative of the population were removed from the experiment, to be used to calculate succulence. For the remaining 21 plants, 11 were transferred to drought conditions, receiving no water, whilst the other 10 stayed in well-watered conditions. Soil moisture readings were taken every two weeks from each pot and results were averaged for each line. Three plants from each genotype were also removed every two weeks and used to calculate succulence.

Succulence was calculated by detaching each part of the above ground biomass and scanning in the images. Using Image J (Version 1.48), the total area of the above-ground biomass could be calculated. Subsequently, all of the above ground biomass was weighed using a balance (Denver Instrument, XP-300). The above ground biomass of each line was then bagged and placed at 70˚C to dry. Once a steady weight was reached, the dry weight of the above-ground biomass was measured which allowed the following calculations to be made which provide insights into leaf succulence:

Succulence: **Leaf fresh weight (FW) / Leaf dry weight (DW)**

Specific Leaf Area (SLA): **Area / Leaf DW**

Dry Leaf Matter Contenet (DLMC): **Leaf DW / Leaf FW**

Avg. Leaf Thickness: **1 / SLA * DLMC**

2.12 Assessing *K. fedtschenkoi* reproductive output

Plants were grown in greenhouse conditions under 16:8 light: dark conditions. Plants were grown until at least leaf pair 14 was fully expanded. Leaves between LP6 and 10 were removed from each plant and placed in trays in the greenhouse to allow the induction of adventitious plantlets. After 14 days, the leaves were examined and the number of leaf margin notches per leaf were counted, as were the number of plantlets starting to form in the leaf notches. From this, reproductive output was calculated as the average number of plantlets produced per leaf notch.

2.13 Bioinformatic Techniques

2.13.1 Identification of genes involved in the short-day dependent induction of flowering and floral meristem development .

Genes that have been defined as playing a role in the day-length dependent induction of flowering and floral meristem development in Arabidopsis and Chrysanthemum were selected to investigate whether orthologous genes are subjected to similar pattrns of regulation in *K. fedtschenkoi*. The amino acid sequence for each flowering gene from Arabidopsis (accessed via TAIR: www.arabidopsis.org), and/ or the translated mRNA sequence from *Chrysanthemum,* were used as query sequences for TBLASTX searches against the draft assembly of the *K. fedtschenkoi* genome using the Custom BLAST feature of Geneious 4.5.5. The assembled genomic scaffold that had the highest E-value, Bit Score and % identical residues, was selected as the likely best ortholog of the query sequence. The region of the genomic scaffold sequence that contained the best BLAST hit to the Arabidopsis or *Chrysanthemum* flowering gene

was used as a query sequence in a Megablast search against the assembled *K. fedtschenkoi* transcriptome. Hits were then assessed again on E-value, Bit Score, % identity, and also, if both an *Arabidopsis* gene and a *Chrysanthemum* mRNA were used, then the results were compared to choose the best hit. This *K. fedtschenkoi* hit was then used as a query sequence in a BLASTP search against the TAIR10 proteins, to check that the putative *K. fedtschenkoi* flowering gene hit the corresponding gene in the Arabidopsis genome as its best BLAST hit (Table 2.9 & 2.10).

Table 2.9. Results gained from TBLASTX searches performed using Geneious 4.5.5 using sequences from the TAIR database

Flowering Gene	Abbreviation	TAIR Accession Number	Scaffold Hit	E-value	Bit Score	% identical sites	Contig Hit	E-value	Bit Score	% identical sites
APETALA1	AP1	AT1G69120	scf1370213	3.63e -55	141.279	95.2	19640	1.21e -32	120.939	88.5
CONSTANS	CO2581	AT5G15840	scf1365388	3.56e -45	148.61	51.6	2581	1.14e -57	182.57	51.6
	C04615						4615	2.56e -42	141.35	100
FD	FD	AT4G35900	scf1352730	1.15e -14	56.1	64.7	11180	1.88e -11	54.7	96.4
FLOWERING LOCUS T	FT12903	AT1G65480	scf1351727	3.48e -73	161.9	82.2	12903	3.60e -27	98.6	70.1
	FT18180	AT5G03840	-	5e-05	205.7	63.2	18180	1.01e -12	58.9	52
FRUITFUL	FUL	AT5G60910	scf1370798	1.07e -49	140.82	93.7	10609	1.27e -32	114.005	93.7
GIGANTEA	GI	AT1G22770	scf1375075	0	351.139	96.8	9732	6.83e -125	375.17	100
SEPALLATA	SEP	AT1G24260	scf1363884	1.07e -55	142.196	96.8	11774	7.46e -03	115.931	60.9
SUPRESSER OF OVEREXPRESSION OF CO1	SOC1	AT2G45660	scf1351697	9.88e -125	81.3	83.1	7510	4.57e -41	138.658	100
ZEITLUPE	ZTL	AT5G57360	scf1359300	0	665.01	83.3	185	0	190.27	96.8

Table 2.10. Results gained from using Geneious 4.5.5 using sequences from the DDBJ database

Flowering Gene	Abbreviation	DDBJ Accession Number	Scaffold Hit	E-value	Bit Score	% identical sites	Contig Hit	E-value	Bit Score	% identical sites
FLOWERING LOCUS T	FT	AB839766	scf1369477	1.51e -44	98.2	56.7	12903	1.33e -23	88.967	63.2
	FT	AB839767	scf1357861	7.34e -53	125.242	74.2	18180	3.39e -24	88.1965	63.1
FD	FD	AB839768	scf1352730	3.26e -02	30.4165	56.5	11180	1.56e -08	45.82	100
	FD	AB839769	scf1374565	1.23e -07	50.0956	65.7	11180	1.07e -03	33.113	66.7
FRUITFUL	FUL	AB839770	scf1354824	1.01e -46	125.97	90.3	10609	9.20e -30	106.301	87.1
APETALLA1	AP1	AB839770	scf1354824	9.20e -30	106.301	87.1	19640	6.97e -25	93.9745	78.7

Chapter 3 – Identification of transgenic lines of *Kalanchoë fedtschenkoi* expressing RNAi or over-expression constructs for the CAM-induced and clock-controlled transcription factors *CAM-INDUCED bZIP1*, *CYCLING DOF FACTOR*, and the single MYB repeat/ *REVEILLE* family gene MYB439

3.1 Introduction

Whilst there remain many gaps in current understanding of the core biochemical steps of CAM, the temporal regulators of the known biochemical steps, especially their control by the circadian clock, also need to be elucidated (Yang *et al*., 2015). This knowledge is needed not only to refine and improve fundamental understanding of temporal coordination of CAM control, but also for projects like "CAM Biodesign" which is funded by the US-Department of Energy, which aims to genetically engineering CAM into C_3 species to improve their water use efficiency (Borland *et al.,* 2014; Yang *et al.,* 2015).

Despite the discovery of the clock-controlled gene encoding the PEPC regulating protein kinase, *PPCK,* over 15 years ago, understanding of the signaling pathway(s) underpinning the daily regulation of CAM relative to the light/ dark cycle and in response to the circadian clock has remained limited (Hartwell *et al.,* 1999; Hartwell 2006; Borland *et al.,* 2009). Nothing is yet known about the regulators of *PPCK* that generate rhythms of nocturnal activity of PPCK in response to the circadian clock. There is therefore an urgent and pressing need to identify the regulatory steps that link the central circadian clock to CAM.

The combination of functional genomic and molecular-genetic approaches can be a powerful method for identifying unknown regulators in a cell-signaling pathway. Such approaches are helped by the use of model species, for which genomic and

transcriptomic resources, mutant populations, and/ or efficient stable transformation protocols exist. For the study of the functional genomics of CAM, the common iceplant (*Mesembryanthemum crystallinum)* has been a preferred model species for a number of years. It has a number of beneficial characteristics including the facultative nature of CAM in this species; where it can switch from C_3 to CAM in response to salt or drought stress (Adams *et al.,* 1998). It also has a relatively small genome, and there are a number of available community resources, including databases for expressed sequence tags, and RNA-Seq gene expression, plus mutant populations (Cushman *et al.,* 2008b; Tsukagoshi *et al.,* 2015). It was also in *M. crystallinum* where the gene encoding the transcription factor CIB1, which is one of the genes manipulated in *K. fedtschenkoi* in this PhD, was first discovered as a CAM-induced, clock-controlled gene. A key factor that has held back the wider adoption of *M. crystallinum* as a model for the study of CAM has been the lack of an efficient system for its stable genetic transformation. Furthermore, the induction of CAM by either salt or water stress introduces an extra layer of complexity to attempts to dissect out the genes and proteins that underpin the transition from C_3 to CAM (Winter *et al.,* 2007; Cushman *et al.,* 2008b).

These limitations have led to the need for other model CAM species. Of these, *K. fedtschenkoi* is popular due to the ease of stable transformation, which allows gene function to be tested *in planta*. *K. fedtschenkoi* is also a constitutive CAM species that shows a clear developmental progression from C_3 in LP1 to full CAM in LP6 (Jones, 1975). Such constitutive CAM species are widely regarded as a 'cleaner' model system for the study of CAM, because they do not require either salt or drought stress in order to perform CAM (Borland *et al.,* 2009). An additional advantage of *K.*

fedtschenkoi is that it can be clonally propagated using leaf margin plantlets, allowing the generation of large experimental populations of identical, developmentally-synchronised plants. *K. fedtschenkoi* has a small genome of 246 Mb, and a draft genome sequence is now available (http://www.cgr.liv.ac.uk/gview/kalanchoe), as well as a number of valuable quantitative RNA-seq datasets which compare C_3 and CAM leaves sampled over the LD cycle, and also epidermal peels and mesophyll tissue from CAM leaves sampled over the light/ dark cycle.

The genus *Kalanchoë* includes a wide diversity of CAM species and has a strong back-catalogue of published CAM-related research. Several key breakthroughs in CAM research were made using *Kalanchoë,* including the discovery of a persistent circadian rhythm of CO_2 fixation associated with CAM (Wilkins 1959; Luttge *et al.,* 1992; Wilkins 1992; Kluge *et al.,* 1996; Luttge 2002), the first clock-controlled CAM-associated regulatory gene; *PPCK* (Hartwell *et al.,* 1996; Hartwell *et al.,* 1999), and more recently, the first published report using transgenic *K. fedtschenkoi* to determine the *in planta* function of two key CAM decarboxylation pathway genes has been published (Chang 1981; Nimmo *et al.,* 1986; Dever *et al.,* 2015).

Although transformation is relatively simple and efficient in *K. fedtschenkoi*, it is highly unlikely to be 100% efficient. Potential factors that could affect the viability of the resulting transgenic lines include insertion of the transgene into an untranscribed region of the genome, insertional mutagenesis of a gene that causes plant death, or the fact that transgene over-expression or knockdown of the endogenous target gene may be lethal to the plant. These and a number of other factors could potentially cause variability in transgene expression across a population of independent primary

transformants. It was therefore important to screen the primary transformants for relevant phenotypes in order to identify the transgenic lines that contained a functioning transgene. In this project, an initial, high-throughput screening pipeline was used in order to identify the best transgenic lines for more detailed analysis.

The leaf punch staining screens employed as a pre-screen here have already been used successfully to screen mutant populations of the CAM plant *M. crystallinum*, allowing the identification of mutant lines which did not have the expected variation in dawn to dusk cell sap pH and leaf starch turnover (Caspar *et al.,* 1985; Cushman *et al.,* 2008a). The basic premise for the pre-screens was that leaf cell sap pH would be low at dawn in a plant performing efficient CAM due to dark period atmospheric CO_2 for re-fixation via PEPC leading to vacuolar malic acid accumulation, causing the chlorophenol red dye to turn yellow. Malate decarboxylation in the light supplying secondary CO_2 fixation would result in a more neutral pH of the cell sap at the end of the day leading to the dye turning red. Equally, leaf starch would accumulate as the product of both photosynthesis and pyruvate recycling in the light period resulting in black stained leaf discs from the iodine solution, whilst leaf starch would reach a minimum at dawn due to its degradation for nocturnal PEP provision for CAM, plus the supply of carbohydrates for nocturnal respiration, and so producing pale leaf discs (Cushman et al., 2008a). The same high-throughput leaf punch staining techniques were used previously in this lab to screen transgenic lines of *K. fedtschenkoi* in which the decarboxylation module genes, *KfNAD-ME_b1* and *KfPPDK* were silenced using RNAi constructs (Dever *et al*., 2015).

These leaf staining screens provided only crude estimations of malate and starch accumulation and turnover that were unlikely to show more subtle perturbations of CAM between the wild type and transgenic lines, but allows identification of lines with clear changes in metabolism, as shown previously (Cushman *et al.,* 2008a; Dever *et al.,* 2015). All lines were also screened for the abundance of the target transcript using semi-quantitative RT-PCR. RNAi lines were expected to show a reduction in the level of the target gene relative to the wild type, whilst the over-expresser lines were expected to express high levels of the full length open reading frame throughout the light/ dark cycle. Combining the results from the acidity and starch screening plus the transcript level of the target gene, it was possible to select transgenic lines with the greatest perturbations in both the transcript abundance of the target transcription factors, and key phenotypes associated with CAM. However, it is important to emphasise that if the transcription factors hypothesized to play a role in the circadian clock control of CAM that were targeted in this work are not in fact important regulators in the light/ dark and circadian control of CAM, then manipulating their transcript levels in transgenic lines would not be expected to have knock-on consequences for phenotypes associated with CAM, such as malate and starch accumulation and turnover. However, as the main focus here was the discovery of novel clock-controlled transcription factors that function to link the central circadian clock to the temporal control of CAM biochemistry, it was valid to limit the selection of the most relevant lines to phenotypes that are directly associated with CAM.

3.2. Results

3.2.1 Identification of transcription factors with potential roles in the Circadian control of CAM

The transcription factors studied in this project were initially identified during a previous BBSRC-funded transcriptomics, quantitative RNA-seq and whole genome sequencing project carried out in the Hartwell lab at the University of Liverpool (2008 – 2011). A key goal of that project was to use *de novo* sequencing of the transcriptome of the model CAM species *K. fedtschenkoi* in order to identify novel candidate regulatory genes that had the potential to be signaling pathway components involved in coupling the temporal outputs from the central circadian clock to the temporal regulation of CAM biochemistry. Total RNA was isolated from leaf pair 1 (LP1, C_3) and leaf pair 6 (LP6, CAM) of *K. fedtschenkoi*. Leaves were sampled in biological triplicate every 4 h over a 12-h-light/ 12-h-dark cycle, starting 2 h after lights on at 02:00. Full length cDNA was synthesized from each RNA sample and sequenced using the Applied Biosystems SOLiD high-throughput sequencing system. The millions of SOLiD reads produced for each RNA/ cDNA library were mapped to the draft genome sequence for *K. fedtschenkoi*, enabling read abundance values to be computed for over 28,000 genes that were detected as expressed in either the C_3 and/ or CAM leaves. Comparison of mean transcript abundance levels (determined as "Fragments per Kb of gene contig length per million reads" = FPKM values) between LP1 and LP6 allowed the identification of genes associated with development of CAM in LP6. Furthermore, those genes that cycled in transcript abundance over the light/ dark cycle could be identified.

As all of the >31,000 genes annotated in the draft *K. fedtschenkoi* genome sequence were linked to putative functions through the known or predicted functions of the

most similar genes in *Arabidopsis thaliana*, it was possible to use the SOLiD RNA-seq dataset to identify putative transcription factor genes which were induced in LP6 and oscillated over the LD cycle (Fig. 3.1A). Transcription factors that fulfilled these criteria were further screened through the design of gene specific primers and the use of semi-quantitative RT-PCR to study the regulation of the transcript under constant free-running circadian conditions (constant light, constant temperature and constant humidity (LL) in CAM leaves (LP6; Fig. 3.1B). Transcription factor genes that were CAM-induced and under light/dark control in the SOLiD RNA-seq dataset, and that also displayed robust circadian oscillations in transcript abundance in the LL experiment were selected for further study using transgenic approaches in *K. fedtschenkoi* (Fig. 3.1). This RNA-seq and LL RT-PCR work was carried out by the researchers employed by the BBSRC *K. fedtschenkoi* sequencing project: Dr. James Hartwell, Prof. Neil Hall, Dr. Susanna Boxall, Dr. Louisa Dever, Dr. Richard Gregory, Jana Knerova and Noemie Meyer.

Once they had identified candidate CAM-induced and clock-controlled transcription factors (Fig. 3.1), it was important to determine whether or not those genes performed a function in the temporal coordination of CAM in response to the central circadian clock. To investigate the *in planta* function of these previously uncharacterized *K. fedtschenkoi* TFs, binary constructs were designed and generated that would either silence the endogenous transcript via the RNAi pathway, through hairpin RNA binary constructs, or over-express the TF constitutively, through use of the CaMV35S promoter to drive the transcription of the full length open reading frame of each identified TF (Dever *et al.,* 2015).

A large number of primary transformants were generated in *K. fedtschenkoi* by Jana Knerova, but the project ended before any detailed screens of the populations could

be carried out, to identify lines that had either the strongest gene silencing phenotype (RNAi lines) or strongest constitutive over-expression phenotype (35S over-expresser lines). This PhD project therefore focused on screening the populations of primary transformants in order to identify the most interesting lines for further, more detailed phenotypic characterisation.

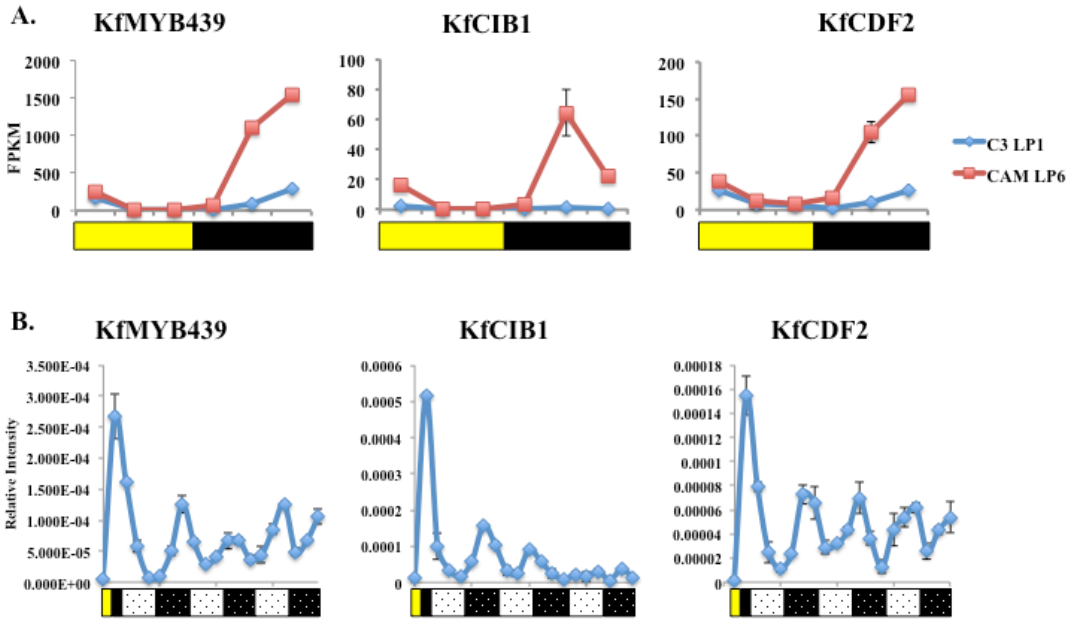

Figure 3.1. Quantitative RNA-seq analysis identified novel, CAM-induced, clock controlled transcription factor genes in *K. fedtschenkoi*. **A.** Diurnal control of transcript abundance (FPKM) for novel CAM-induced genes in leaf pair 1 (LP1, C_3) and leaf pair 6 (LP6, CAM) from SOLiD RNA-seq data. The blue lines represent C_3 leaves (LP1), and the red lines represent CAM leaves (LP6). Experiment conducted by Hartwell lab technician Jane Knerova **B.** Semi-quantitative RT-PCR using RNA from leaves sampled every 4 h in LL showing robust circadian oscillations for many of the novel CAM-induced transcripts. The yellow bars represent lights on and black bars show lights off. Experiment conducted by Jane Knerova.

Figure 3.1 shows that the transcript abundance of TF *KfMYB439* was subject to strong light/ dark regulation in both C_3 and CAM leaves. In both leaf ages, the highest level of this transcript occurred at the end of the dark period at 22:00, 2 h before dawn, and the transcript rose mainly in the second half of the dark period, being high at both 18:00 and 22:00 (Fig. 3.1). Comparing the level of *KfMYB439* between LP6 (CAM)

and LP1 (C_3), the transcript abundance was up to 12-fold greater in LP6 (CAM) at 18:00, and reached a peak FPKM value of 1540 at 22:00, whereas in LP1 the peak FPKM value at 22:00 was only 289 (Fig. 3.1). When RT-PCR was used to monitor the transcript abundance of *KfMYB439* under LL free-running conditions, the transcript was low at the end of the light, and high at the end of the dark in the final LD samples before the plants entered LL (Fig. 3.1B). This matched with the regulation observed in the RNA-seq data (Fig. 3.1A). The transcript level then oscillated robustly over the 3 days in LL, with transcript peaks initially in phase with the 22:00 peak observed under LD, but soon advancing ahead of this, consistent with the well-known short-period free-running rhythm of CO_2 fixation and gene regulation in *K. fedtschenkoi* under LL conditions (Anderson & Wilkins, 1989; Dever *et al*., 2015). The amplitude of the free-running rhythm of *KfMYB439* remained high throughout the 3 days in LL indicating very robust circadian control of this gene in LL conditions (Fig. 3.1B).

KfCIB1 transcript levels also displayed a robust light/ dark oscillation in LP6 in the RNA-seq data, with peak transcript levels being reached at 18:00 in the middle of the dark period (Fig. 3.1A). The FPKM value for *KfCIB1* at its peak at 18:00 in LP6 was 64, whereas the FPKM value in LP1 at this time was 1 (Fig. 3.1B). Thus, *KfCIB1* displayed a >60-fold induction in LP6 relative to LP1 at its time of peak transcript abundance. It is noteworthy that the FPKM value for *KfCIB1* at its peak was 24-fold lower than the FPKM value for *KfMYB439* at its peak. Thus, *KfCIB1* is a much lower abundance transcript than *KfMYB439*. Under LL free-running conditions, RT-PCR revealed that *KfCIB1* transcripts were low at the end of the last light period and high at the end of the last dark period prior to LL(10:00 and 22:00 LD), consistent with the regulation of this gene detected under LD cycles using RNA-seq (Fig. 3.1A & 3.1B).

Once into LL conditions, the oscillation in *KfCIB1* transcript levels dampened rapidly towards arrhythmia, with only two clear, low-amplitude peaks in during LL. This suggests that *KfCIB1* is under only weak circadian clock control, and that it requires daily light/ dark signals in order to maintain high-amplitude oscillations in its transcript abundance.

Finally, *KfCDF2* transcript levels peak at the end of the dark, 22:00, in both tissues, and showed a huge increase of transcript once lights are off, with a 5.6 fold rise between 2h and 6h dark in CAM tissues. In LP1 the peak FPKM value at 22:00 reached 27, whereas in LP6 the value was 155, showing a 5.7 fold induction of transcript in leaves conducting CAM (Fig. 3.1A). RT-PCR was used to investigate transcript regulation in LL constant free-running circadian conditions. This data supports the sequencing results obtained, showing that in the final light and dark sample before entering LL, much greater amounts of transcript were present at 10h dark and barely any present at 10h light (Fig 3.1B). Once in free running conditions *KfCDF2* oscillates robustly for three days, like *KfMYB439*, and also shows the same short-period free-running rhythm of *KfCDF2* expression, indicating this transcript is also robustly regulated by the circadian clock.

From these three TFs, *KfCIB1* is the most highly induced, with a fold increase of 64 at peak expression time at 18:00. *KfCDF2* has a 5.7 fold induction and *KfMYB439* was induced the least with 5.3 fold increase at 22:00. Despite *KfCIB1* being the most induced, the most highly expressed line is *KfMYB439* with a peak FPKM value of 1540 in LP6, whereas *KfCDF2* has an FKPM value of 155 and *KfCIB1* only reaches 64. Therefore if transcript levels reflect corresponding protein levels, it would suggest *KfMYB439* in LP6 is able to interact with a lot more of the genome, possibly including CAM-related genes, than the other two TFs. Although if the other TFs

target a more specific group of (CAM) genes, or just a fewer number, then they may not be required in such large quantities in the leaf.

The rationale for generation of the studied lines was that RNAi silencing would produce phenotypes similar to a knockout mutant, whilst constitutive over-expression might cause arrhythmia in downstream processes controlled by the manipulated TF, especially if any of the targeted TF genes were involved in a circadian feedback loop. This has been demonstrated for *CCA1*-overexpression in *A. thaliana,* which caused arrhythmia of multiple clock outputs (Wang and Tobin, 1998; Dodd *et al*., (2005), and also for CO_2 fixation in *TOC1-OX* lines of *K. fedtschenkoi,* where the over-expression of *TOC1* caused arrhythmia of the CAM circadian rhythm of CO_2 fixation (Dall'omo 2011).

3.2.2 Screening of *K. fedtschenkoi* RNAi and over-expressor lines for *KfMYB439*, *KfCIB1* and *KfCDF2*

3.2.2.1 *KfMYB439*

K. fedtschenkoi gene *KfMYB439* (gene KF79015 in the draft genome assembly) has the highest sequence identity to an *A. thaliana* gene called *HOMEODOMAIN SUPERFAMILY PROTEIN* (At3g09600), which includes a MYB DNA binding domain related to that of the central oscillator genes *CCA1* and *LHY* (Riechmann *et al.,* 2000). Specifically, At3g09600 is a member of the REVEILLE-family of single MYB-repeat transcription factors that are important for the operation of the core molecular circadian oscillator in plants (Hsu *et al.,* 2013). At3g09600 has been designated as *RVE8,* although *KfMYB439* also has high levels of similarity to other TFs from the RVE family; *RVE4* (At5g02840) and *RVE6* (At5g52660). Current evidence suggests that these three TFs, *RVE4*, *6* and *8* contribute to an afternoon phased peak of evening element (EE) binding activity (Hsu *et al.,* 2013). EEs are found in the promoters of a wide diversity of evening-phased, clock-controlled genes, including core clock genes such at *TOC1* (Harmer *et al.,* 2000; Michael *et al.,* 2002). It was therefore reasoned that this transcription factor might be involved in the core mechanism of the *K. fedtschenkoi* circadian clock, like the related *RVE* genes in Arabidopsis. However, whilst *KfMYB439* displayed 5- to 14-fold induction in CAM leaves relative to C_3, the *K. fedtschenkoi* orthologues of *CCA1*/ *LHY* displayed only 2- to 6-fold induction with CAM. In order to further elucidate the function, if any, of *KfMYB439* in the circadian coordination of CAM, both RNAi knock-down and 35S constitutive over-expressor lines had been generated previously by Jana Knerova.

3.2.2.1.1 *KfMYB439* over-expresser lines

For the full length open reading frame over-expression lines for *KfMYB439*, 35 transgenic lines were established in soil at the start of this project. Using RT-PCR, twelve lines were confirmed as over-expressing the full-length transcript at both 2 hours before dusk (10:00) and 2 hours before dawn (22:00; Table 3.1). The primers used to determine the overexpression target both the endogenous transcript and also the transgene, and so it is worth noting that the dark results for these lines are the combination of all *KfMYB439* transcripts present in the sampled leaf at that time. These lines showed large variations in the level of over-expression of the transgene, for example, line *KfMYB439_FL_37C* showed a 38-fold increase in the level of total *KfMYB439* transcript in the light relative to the wild type, but this same line displayed only a 1.1-fold increase in transcript abundance in the dark (Fig. 3.2). By contrast, line *KfMYB439_FL_54B* showed a 24-fold increase in *KfMYB439* transcript at the end of the light period, and at 22:00, peak expression time for *KfMYB439*, the fold change showed an increase of 1.7. Despite the fact that the CaMV35S constitutive promoter was used, the transcript abundance of the *KfMYB439* transgene was not found to be uniformly high in both the light and dark RNA samples, but instead was found to vary in abundance between light and dark, likely due to the endogenous transcript also being amplified by the primers used for analysis. However, the fold-change in the abundance of *KfMYB439* transcript levels between light and dark was dramatically lower for the transgenic lines than the wild type. Wild type shows a transcript increase of 52 fold when comparing the low levels present 2h before dusk to the levels present 2h before dawn. Both 37C and 54B show much less of a change, with 37C showing a 1.2 fold increase in dark transcript level, and 54B showing a 3.5 fold increase from

light to dark. This therefore demonstrates that in these lines *KfMYB439* is being effectively overexpressed.

In addition to screening the lines for the transcript abundance of the transgene using RT-PCR, all lines were also screened for leaf cell-sap pH at dawn and dusk (as a proxy for malate levels resulting from dark CO_2 fixation and light period malate decarboxylation), and also dawn and dusk starch levels. Screening these two metabolite pools revealed that these *KfMYB439* overexpressing lines tended to accumulate less malate than the wild type by the end of the dark period, and have a higher concentration of malate than the wild type remaining at the end of the light period (Table 3.1). This indicates that the lines were performing less CAM. Interestingly, two lines at both time points had no noticeable differences to the WT, and two consistently similar results to WT, namely: 37 and 56C (Table 3.1). At dusk 2A also displays similar levels to WT, but at dawn has higher malate concentrations than WT, whereas *KfMYB439_FL_17* has higher malate concentrations at dusk, but at dawn had similar amounts to WT.

Table 3.1. Summary table for *KfMYB439_FL* screening results for successful overexpressing lines.

Transgene being manipulated	Line	Transcript change relative to WT - Light	Transcript change relative to WT - Dark	pH relative to WT - Dusk	pH relative to WT - Dawn	Starch relative to WT - Dusk	Starch relative to WT - Dawn
KfMYB439 FL	1	17.0	1.3	<	<	=	=
KfMYB439 FL	1A	16.0	1.4	<	<	<	<
KfMYB439 FL	2A	10.1	1.3	=	>	=	=
KfMYB439 FL	17	10.8	1.6	>	=	=	=
KfMYB439 FL	19	9.7	1.5	>	<	=	=
KfMYB439 FL	27A	10.9	1.3	>	<	=	=
KfMYB439 FL	27B	21.7	1.5	>	<	<	<
KfMYB439 FL	33A	15.4	1.5	>	<	=	=
KfMYB439 FL	37	27.5	1.1	=	=	=	=
KfMYB439 FL	37C	38.0	1.1	<	<	<	<
KfMYB439 FL	54B	24.4	1.7	>	<	<	<
KfMYB439 FL	56C	19.0	1.0	=	=	=	=

Results of fold reduction of transcript in relation to WT (10h light and dark), Malate (pH) and starch leaf disc screen at 11h Light and 11h Dark over a 12:12 LD cycles for the confirmed transgenic lines.All results are in comparison to an average wild type of three biological replicates. For pH, WT is less acidic at dusk and highly acidic at dawn, and so symbols represent lines showing higher, lower or equal acidity than wild type at that time point accompanied with either red indicating a decrease, yellow indicating no change and green showing an increase. For starch levels the symbols represent higher, lower or equal starch levels at dawn and dusk. At dawn starch is low in wild type and at dusk there are high levels present. < indicates a lower pH or a lower amount of starch present, > shows a higher value for the transgenic and = represents no difference to the WT.

Fewer differences were detected in leaf starch levels at dawn and dusk in the *KfMYB439_FL* lines using the iodine staining test to estimate starch levels in punched leaf discs. It is probable that few lines were observed to be different to the wild type as this assay is not quantitative, a being purely visual test that is realistically only able to discern presence/ absence of starch in the leaf. Despite this limitation, four lines; 1A, 27B, 37C and 54B were found to have consistently lower starch both at the end of the light and the end of the dark (Table 3.1).

From these initial pre-screens of the entire population of putative *KfMYB439_FL* lines, two lines were selected for further analysis, namely *KfMYB439_FL_37C* and *KfMYB439_FL_54B* (Fig. 3.2; Table 3.2). These lines were chosen on the basis that line 37C was the highest over-expresser in the light and line 54B was an intermediate over-expresser. Line 54B also had the highest level of dark over-expression of *KfMYB439* transcripts, whereas line 37C showed, at best, only a very small degree (1.1-fold) of over-expression in the dark (Fig. 3.2). Furthermore, both lines shows changes in malate and starch levels over the 24h cycle, suggesting proper functioning of CAM may have been perturbed (Table 3.2).

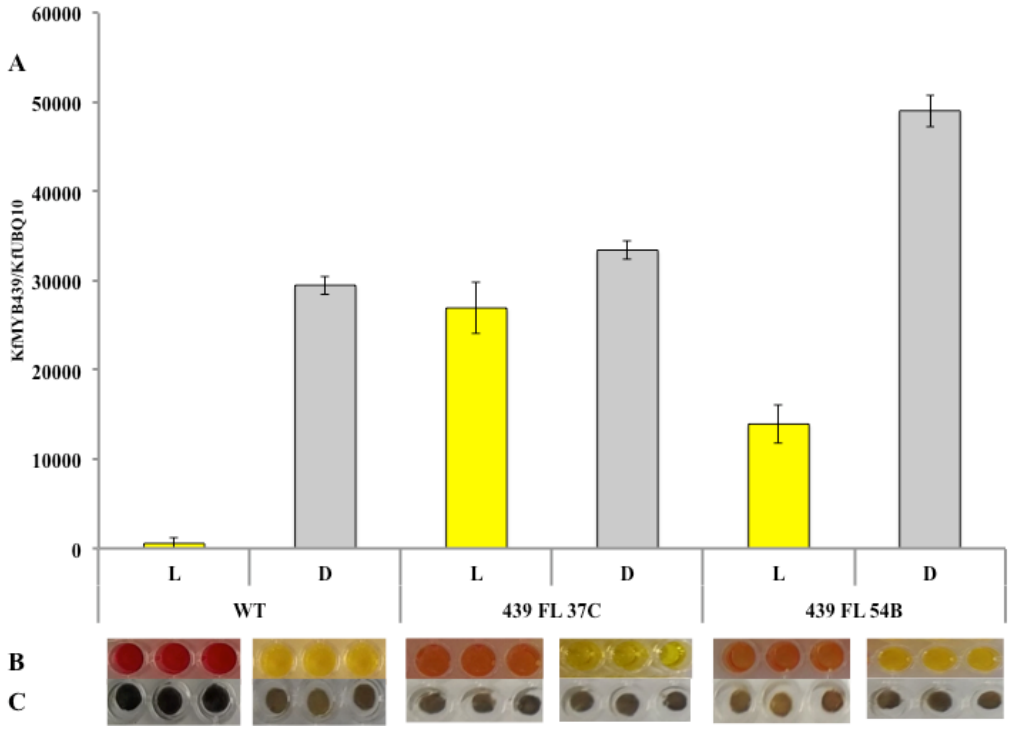

Figure 3.2. Transcript level, malate and starch results of *KfMYB439_FL* in Wild Type and the two chosen lines to be used for further analysis. **A.** Showing transcript level is naturally very low during the day, and peaks during the night. The overexpressing lines; 37C and 54B, show a drastic increase in light expression. And a smaller, but still increase, during the dark. **B.** Shows chlorophenol red pH indicator dye after leaf discs have been incubated and removed. Red indicates a pH of 6 or over, whereas yellow indicates pH4 or below, indicating high levels of malate are present. **C.** Levels of starch over the 24 hour cycle. At the end of light there is much starch produced in Wild Type from photosynthesis and so leaves appear black in the iodine stain. At the end of the dark, starch has been used for PEP synthesis, and so lighter coloured discs should be seen.

Table 3.2. Summary screening results for the two *KfMYB439_FL* over expressing lines being taken forward for further, more detailed analysis.

	MYB439 FL 37C	MYB439 FL 54B
Fold increase of gene expression - Light	38.0	24.4
Fold increase of gene expression - Dark	1.1	1.7
Malate - Light	Low	High
Malate - Dark	Low	High
Starch – Light	Low	Low
Starch - Dark	Low	Low

3.2.2.1.2 *KfMYB439* RNAi

For the RNAi knockdown lines for *KfMYB439*, only 8 plants were established on soil to screen. This is because high levels of plantlet death occurred. These 8 plants also grew very slowly compared to other wild type and transgenic lines. Each line was screened at 2h before dawn (22:00), as this is when wild type transcript levels are highest, so would be most noticeable in the RNAi lines that transcript reduction was occurring. Interestingly, when RT-PCR was conducted, all 8 plants were confirmed to have a reduction in transcript, averaging 1.3 fold less than WT (Table 3.3). The highest reduction in transcript was 1.5-fold, which is interesting as this could indicate why only 8 plants survived, as if this transcript was important to plant metabolism, there may be a minimum threshold level needed to sustain life. This could also explain why all 8 lines showed very stunted growth, if they are already under high stress due to KfMYB439 protein reduction.

Table 3.3. Summary table for *KfMYB439_RNAi* screening results

Transgene being manipulated	Line	Transcript change relative to WT - Dark	pH relative to WT - Dusk	pH relative to WT - Dawn	Starch relative to WT - Dusk	Starch relative to WT - Dawn
KfMYB439 RNAi	2A	1.2	=	>	<	=
KfMYB439 RNAi	2B	1.2	=	>	=	=
KfMYB439 RNAi	3C	1.5	>	<	<	=
KfMYB439 RNAi	4B	1.1	=	=	=	=
KfMYB439 RNAi	5	1.3	<	<	<	<
KfMYB439 RNAi	9A	1.3	>	>	=	=
KfMYB439 RNAi	10B	1.3	>	>	<	<
KfMYB439 RNAi	12	1.4	>	=	<	<

Results of fold reduction of transcript in relation to WT (10h light and dark), Malate (pH) and starch leaf disc screen at 11h Light and 11h Dark over a 12:12 LD cycles for the confirmed transgenic lines.All results are in comparison to an average wild type of three biological replicates. For pH, WT is less acidic at dusk and highly acidic at dawn, and so symbols represent lines showing higher, lower or equal acidity than wild type at that time point accompanied with either red indicating a decrease, yellow indicating no change and green showing an increase. For starch levels the symbols represent higher, lower or equal starch levels at dawn and dusk. At dawn starch is low in wild type and at dusk there are high levels present. < indicates a lower pH or a lower amount of starch present, > shows a higher value for the transgenic and = represents no difference to the WT.

The 8 lines were screened for malate and starch levels, which showed a large range of phenotypes (Table 3.3). The line which is the lowest knockdown line, *KfMYB439_RNAi_4B* with only a .1 increase, showed no differences from the wild type. As the level of transcript reduction increases, more changes to the malate and starch results can be seen.

After this initial screening was completed, two lines were selected for further screening, again with varying differences in the amount of transcript reduction; *KfMYB439_RNAi_3C* and *KfMYB_RNAi_10B* (Fig. 3.3; Table 3.4).

KfMYB439_RNAi_3C showed the highest reduction in transcript with a 1.5 fold reduction, which also appears to have had knock-on effects to both malate and starch pools. *KfMYB439_RNAi_10B* showed an intermediate level of reduction of 1.3 fold,

which also began to show differences to both malate and starch metabolite pools over the 24h cycle (Fig. 3.3, Table 3.4). For both lines this therefore indicates CAM performance is no longer optimal when levels of *KfMYB439* are reduced, even by small amounts.

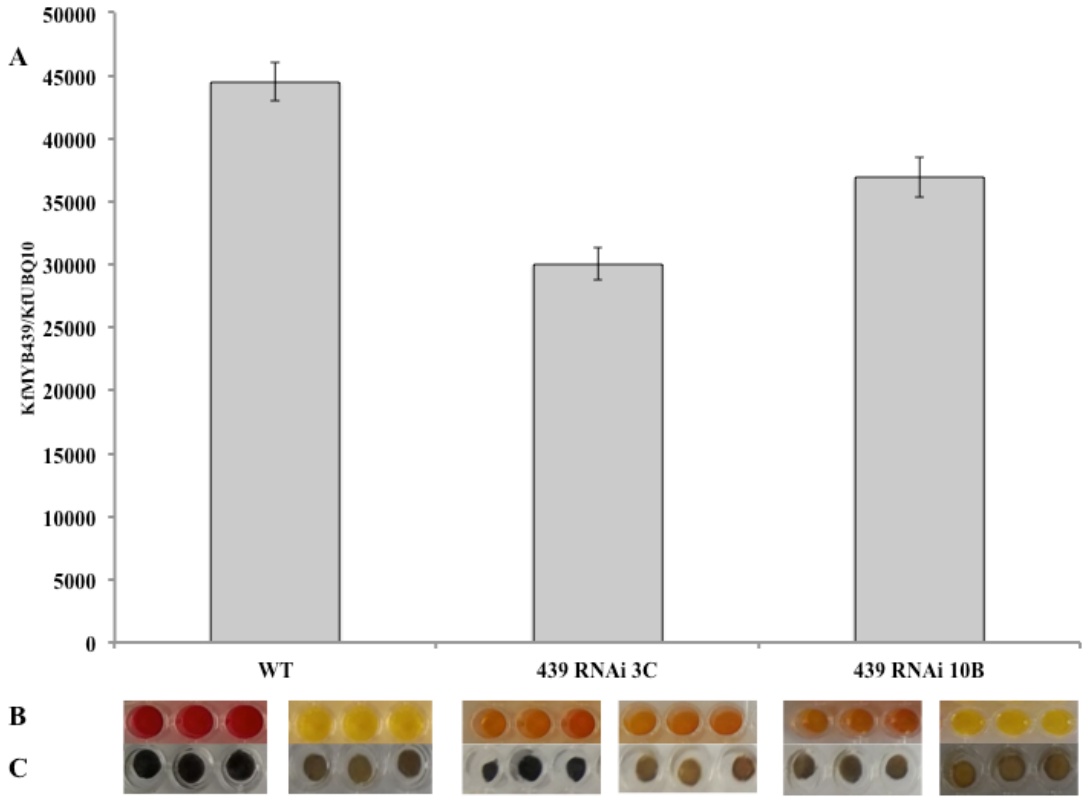

Figure 3.3. Transcript level, malate and starch results of *KfMYB439_RNAi* in Wild Type and the two chosen lines to be used for further analysis. **A.** Showing RT-PCR results collected at 22:00h (2h before dawn), as this is when *KfMYB439* peaks in WT. Both *KfMYB439_RNAi_3C* and 10B show reduction in levels, with *KfMYB439_RNAi_3C* showing the biggest reduction. **B.** Shows chlorophenol red pH indicator dye after leaf discs have been incubated and removed. End of light screening is on the left hand side and dark is on the right. Red indicates a pH of 6 or over, whereas yellow indicates pH4 or below, indicating high levels of malate are present. For WT a red colour is expected at the end of light, due to malate decarboxylation during the day, and the dye is yellow at the end of the dark due to malic acid accumulation. **C.** Levels of starch at the end of light (left) and end of dark (right). WT has high starch levels at the end of the light (dark discs) and low levels at the end of the dark (pale discs).

Table 3.4. Summarising screening results for the two *MYB439_FL* over expressers being taken forward for further analysis.

	MYB439 RNAi 3C	MYB439 RNAi 10B
Fold decrease of gene expression - Dark	1.5	1.3
Malate - Light	High	High
Malate - Dark	Low	High
Starch – Light	Low	Low
Starch - Dark	No Visible Difference	Low

3.2.2.2 *KfCIB1*

The *KfCIB1* sequence (gene KF13220 in the draft genome assembly has shown highest identity with a BASIC LEUCINE ZIPPER (bZIP) transcription factor (At3g30530). The bZIP family is one of the largest and most functionally diverse families found in plants, making up 4% of all Arabidopsis TFs (Bartels *et al.,* 2005). There are many groups within the bZIP superfamily, and *KfCIB1*s ortholog is located within group S (Meszter 2010). This group of genes are known to be involved in adaptation to environmental factors and control of energy use (Correa *et al.,* 2008). Therefore as CAM is an extreme environmental adaption, and it has been seen to show a circadian rhythm, it was reasoned this transcription factor could play a key role in linking the central oscillator to CAM. This seems increasingly likely from the SOLiD data, where *KfCIB1* shows between 32- to 61-fold inductions in transcript in LP6 relative to LP1. Overexpression and RNAi knockdown lines were created.

3.2.2.2.1 *KfCIB1* over-expresser lines

For the *KfCIB1_FL* lines, 42 lines were generated. Of these, after transcript level screening using semi-quantitative RT-PCR, 34 were confirmed as overexpressing. WT transcript levels are low at 10:00 (2 hours before dusk) and are high at 22:00 (2 hours before dawn). From the transcript level screening, overexpression levels were quite high, with an average fold transcript increase of 2.1 during the light period and 1.7 during the dark period. Again it is worth noting that the primers used to assess transcript abundance with RT-PCR amplified both the endogenous and the transgene, and so results are an accumulation of the two. All lines showed a large increase in transcript levels during the light (10:00), when WT is relatively low, and many lines also showed large dark increases too. The CaMV35S promoter is clearly working

effectively as the average fold change for wild type between dusk and dawn (10:00 and 22:00) transcript level is 2.6, whereas with *KfCIB1_FL* lines, such as 5C the fold change is much lower at 0.7, as the transcript level is now higher in the light than the dark, and with *KfCIB1_FL_11B* the fold change is now only 1.1, showing constant high overexpression (Fig. 3.4).

Malate (sap pH) and starch screens were also carried out on the plants that were confirmed to overexpress the transgene (Table 3.5). From the malate assay, there appeared to be a trend of reduced malic acid present in these transgenic lines compared to wild type, with 17 lines showing lower levels at dusk and 20 lines having lower levels at the end of the dark period (Table 3.5). Due again to the starch being a visual screen the majority of the lines showed no noticeable difference to wild type, although seeing an increase in starch levels at dusk would be especially difficult, as the wild type would already be stained a dark black from the iodine, and so there would be no way to be able to differentiate between the wild type and transgenic lines. From the lines that did show a difference, most showed a lower amount of starch, with 6 lines at dusk and 10 at dawn (Table 3.5).

Table 3.5. Summary table for *KfCIB1_FL* screening results of successful overexpressing lines.

Transgene being manipulated	Line	Transcript change relative to WT - Light	Transcript change relative to WT - Dark	pH relative to WT - Dusk	pH relative to WT - Dawn	Starch relative to WT - Dusk	Starch relative to WT - Dawn
KfCIB1 FL	3A	2.8	2.5	<	=	=	=
KfCIB1 FL	3B	2.5	2.1	<	>	=	>
KfCIB1 FL	3C	1.7	1.3	<	=	=	=
KfCIB1 FL	5B	6.4	1.6	<	<	<	=
KfCIB1 FL	9B	2.1	1.9	=	=	=	=
KfCIB1 FL	10A	2.8	2.2	>	<	=	<
KfCIB1 FL	10B	1.6	1.6	>	<	=	<
KfCIB1 FL	10BB	2.5	2.3	<	<	=	<
KfCIB1 FL	10C	2.8	2.4	<	<	=	=
KfCIB1 FL	11A	1.1	1.4	=	=	=	=
KfCIB1 FL	11B	3.1	1.3	>	>	=	=
KfCIB1 FL	12A	2.2	1.9	<	<	<	=
KfCIB1 FL	12B	2.5	2.2	=	=	=	=
KfCIB1 FL	12C	1.2	1.4	>	<	=	=
KfCIB1 FL	16B	1.8	1.6	>	>	=	>
KfCIB1 FL	16C	1.6	1.7	>	>	=	>
KfCIB1 FL	17	1.3	1.4	<	<	=	=
KfCIB1 FL	17A	1.7	1.4	<	<	<	=
KfCIB1 FL	17B	2.1	1.8	<	=	=	<
KfCIB1 FL	17C	2.1	1.9	<	<	=	<
KfCIB1 FL	19B	1.5	1.6	>	<	=	<
KfCIB1 FL	23B	1.8	1.7	>	<	=	>
KfCIB1 FL	23C	1.7	1.3	<	<	=	=
KfCIB1 FL	24A	1.7	1.5	<	<	<	<
KfCIB1 FL	24B	1.9	1.6	<	<	=	>
KfCIB1 FL	24C	1.6	2.1	<	<	<	<
KfCIB1 FL	25A	1.5	1.3	=	=	=	=
KfCIB1 FL	25B	1.9	1.7	=	=	=	=
KfCIB1 FL	25C	2.3	2.3	=	=	=	<
KfCIB1 FL	26B	1.9	1.7	<	<	=	=
KfCIB1 FL	26C	1.3	1.3	=	<	=	=
KfCIB1 FL	28A	2.2	2.1	=	=	=	=
KfCIB1 FL	28C	1.8	1.7	<	<	=	<
KfCIB1 FL	31A	1.7	1.5	=	<	<	=

Results of fold reduction of transcript in relation to WT (10h light and dark), Malate (pH) and starch leaf disc screen at 11h Light and 11h Dark over a 12:12 LD cycles for the confirmed transgenic lines.All results are in comparison to an average wild type of three biological replicates. For pH, WT is less acidic at dusk and highly acidic at

dawn, and so symbols represent lines showing higher, lower or equal acidity than wild type at that time point accompanied with either red indicating a decrease, yellow indicating no change and green showing an increase. For starch levels the symbols represent higher, lower or equal starch levels at dawn and dusk. At dawn starch is low in wild type and at dusk there are high levels present. < indicates a lower pH or a lower amount of starch present, > shows a higher value for the transgenic and = represents no difference to the WT.

After these screens two out of the 36 over expressers were picked to be carried forward for more detailed analysis on how CAM was perturbed; *KfCIB1_FL_5B* and *KfCIB1_FL_11B* (Fig. 3.4; Table 3.6). *KfCIB_FL_5B* was chosen because it was the highest overexpressing line during the light, and was also interesting because there appeared to be much greater amounts of *KfCIB1* transcript present at dusk rather than dawn, which may have interesting consequences in metabolism. This line also demonstrated low levels of malate at both time points, and also showed a decrease in starch at dusk. *KfCIB1_FL_11B* was chosen due to an intermediate level of overexpression, and also due to the level being very similar at both dusk and dawn, suggesting a possible change in circadian control. It was also interesting that this level of expression appeared to lead to increases in malate at dusk and dawn, whereas higher overexpression appears to lead to reductions, as seen in *KfCIB1_FL_5B*.

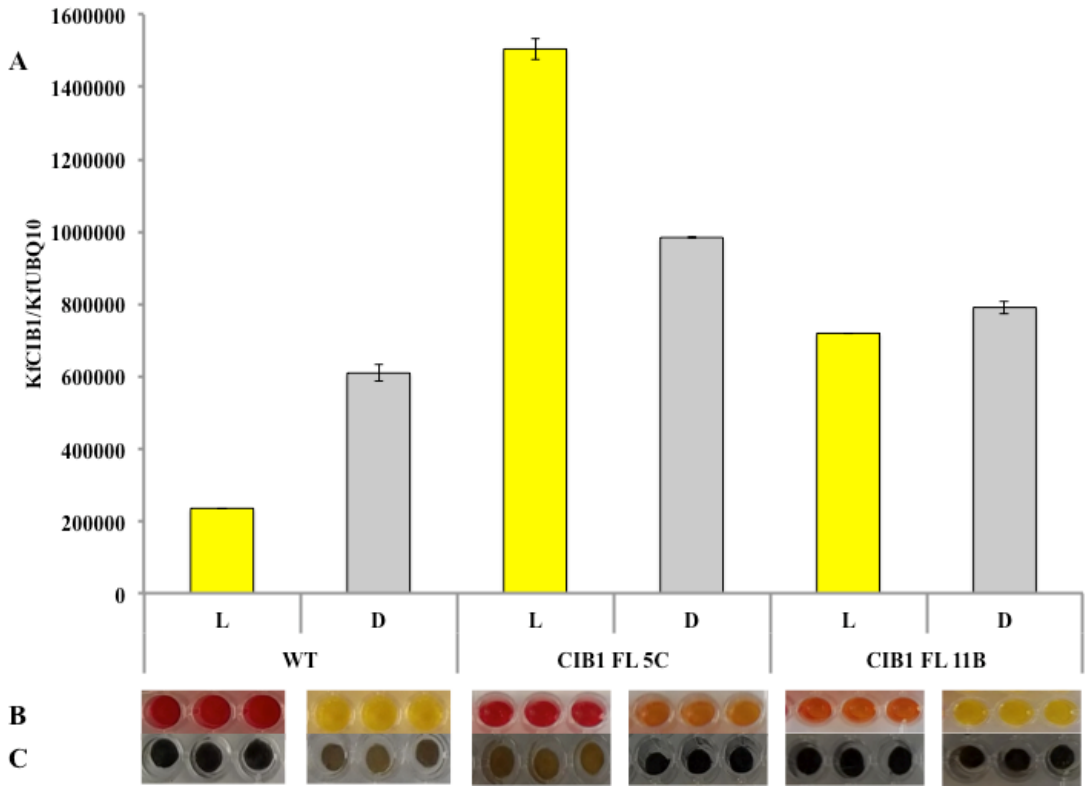

Figure 3.4. Transcript level, malate and starch results of KfCIB1 in Wild Type and the two chosen lines to be used for further analysis. **A.** Showing transcript naturally peaks during the night. The overexpressing lines; 5C and 11B, show a drastic increase in light expression. And a smaller increase during the dark. **B.** Shows chlorophenol red pH indicator dye after leaf exposure. Red indicates a pH of 6 or over, whereas yellow indicates pH4 or below. **C.** Levels of starch over the light/dark cycle. Dark leaf discs indicate high starch and pale leaf discs show low start.

Table 3.6. Summarising screening results for the two *CIB1_FL* over expressing lines being taken forward for further analysis.

	CIB1 FL 5C	CIB1 FL 11B
Fold increase of gene expression - Light	6.4	3.1
Fold increase of gene expression - Dark	1.6	1.3
Malate - Light	Low	High
Malate - Dark	Low	High
Starch – Light	Low	No Visible Difference
Starch - Dark	No Visible Difference	No Visible Difference

3.2.2.2.2 *KfCIB1 RNAi* lines

There were 67 lines generated from the *KfCIB1 RNAi* construct. Out of these 50 were confirmed as having a reduction in transcript level relative to the Wild Types, with an average fold reduction of 2.3 (Table 3.7). Lines were entrained to 12:12 LD conditions for 7 days before leaf pair 6 was sampled at 22:00, 2h before dusk, when endogenous *KfCIB1* transcript levels are still raised.

The 50 knockdown lines were screened for malate (sap pH) and starch changes relative to wild type. For malate levels 14/50 lines showed no difference to wild type at either dawn or dusk, with 12 of these lines showing a less than 2 fold reduction in transcript, suggesting larger decreases are required to impact strongly on malate metabolite pool accumulation (Table 3.7). With the starch assays, again most lines were found to be similar to wild type, with 27 lines showing no difference at dawn or dusk. 10 lines have been shown to have no differences to wild type in either the malate and starch screen, which again may be due to low transcript reduction, with transcript reductions of these lines ranging from 1.1 to 1.9 fold.

From this, two lines were chosen to bring forward for further, detailed investigation; *KfCIB1_RNAi_12A* and *KfCIB_RNAi_30A* (Fig. 3.5; Table 3.8).

Table 3.7. Summary table for *KfCIB1_RNAi* screening results

Transgene being manipulated	Line	Transcript change relative to WT - Dark	pH relative to WT - Dusk	pH relative to WT - Dawn	Starch relative to WT - Dusk	Starch relative to WT - Dawn
KfCIB1 RNAi	1B	1.7	>	<	<	<
KfCIB1 RNAi	2A	1.6	>	<	=	<
KfCIB1 RNAi	2B	1.7	>	<	=	<
KfCIB1 RNAi	4A	1.2	>	<	<	<
KfCIB1 RNAi	7A	1.8	=	=	=	=
KfCIB1 RNAi	7B	1.4	<	=	=	=
KfCIB1 RNAi	7C	1.9	<	>	>	<
KfCIB1 RNAi	8A	1.7	=	<	=	>
KfCIB1 RNAi	9B	2.7	>	=	=	=
KfCIB1 RNAi	9C	2.2	=	=	<	=
KfCIB1 RNAi	10A	2.1	=	<	=	<
KfCIB1 RNAi	11A	1.3	=	=	=	=
KfCIB1 RNAi	12A	1.7	>	=	=	<
KfCIB1 RNAi	12B	1.3	=	=	=	=
KfCIB1 RNAi	12C	4.1	>	=	<	<
KfCIB1 RNAi	13A	1.1	<	<	<	<
KfCIB1 RNAi	13C	1.1	=	=	=	=
KfCIB1 RNAi	14C	1.2	<	<	=	=
KfCIB1 RNAi	15B	1.3	=	=	=	=
KfCIB1 RNAi	16B	1.3	<	=	<	=
KfCIB1 RNAi	19A	1.3	=	=	=	=
KfCIB1 RNAi	20A	1.1	=	<	=	=
KfCIB1 RNAi	20B	1.1	=	<	=	=
KfCIB1 RNAi	21B	1.2	=	=	=	=
KfCIB1 RNAi	22A	1.6	=	=	=	=
KfCIB1 RNAi	23B	1.1	=	<	=	=
KfCIB1 RNAi	24A	1.2	<	<	=	<
KfCIB1 RNAi	26B	1.5	<	<	<	<
KfCIB1 RNAi	26C	1.7	=	>	=	<
KfCIB1 RNAi	28A	1.9	=	=	=	=
KfCIB1 RNAi	28B	1.9	=	<	=	=
KfCIB1 RNAi	28C	5.4	<	<	=	=
KfCIB1 RNAi	29B	3.2	<	<	=	=
KfCIB1 RNAi	29C	5.1	<	=	=	=
KfCIB1 RNAi	30A	10.8	<	<	<	=
KfCIB1 RNAi	30B	7.9	<	=	=	=
KfCIB1 RNAi	30C	4.7	>	=	=	=
KfCIB1 RNAi	36C	1.4	<	=	<	<
KfCIB1 RNAi	41B	1.6	<	=	=	=

KfCIB1 RNAi	41C	1.5	=	<	=	=
KfCIB1 RNAi	42B	1.6	<	<	=	<
KfCIB1 RNAi	49A	1.2	<	<	<	<
KfCIB1 RNAi	49C	1.3	=	>	=	<
KfCIB1 RNAi	50A	1.5	=	=	=	=
KfCIB1 RNAi	52B	1.7	=	<	=	=
KfCIB1 RNAi	52C	1.1	<	<	=	=
KfCIB1 RNAi	55C	1.6	<	<	=	>
KfCIB1 RNAi	56A	1.2	=	=	=	=
KfCIB1 RNAi	56C	1.5	=	=	<	=
KfCIB1 RNAi	57A	1.2	=	<	=	<

Results of fold reduction of transcript in relation to WT (10h light and dark), Malate (pH) and starch leaf disc screen at 11h Light and 11h Dark over a 12:12 LD cycles for the confirmed transgenic lines.All results are in comparison to an average wild type of three biological replicates. For pH, WT is less acidic at dusk and highly acidic at dawn, and so symbols represent lines showing higher, lower or equal acidity than wild type at that time point accompanied with either red indicating a decrease, yellow indicating no change and green showing an increase. For starch levels the symbols represent higher, lower or equal starch levels at dawn and dusk. At dawn starch is low in wild type and at dusk there are high levels present. < indicates a lower pH or a lower amount of starch present, > shows a higher value for the transgenic and = represents no difference to the WT.

CIB1_RNAis 12A and *30A* were picked, showing 1.7 and 10.8 fold knockdown respectively (Fig. 3.5). Both also had malate and starch results that differed from the wild type and also each other (Table 3.8). Although the results for the two lines do not have many results which are the same, it was decided it was still best to look at an intermediate knockdown and a high knockdown, to be able to investigate the dosage effects on CAM.

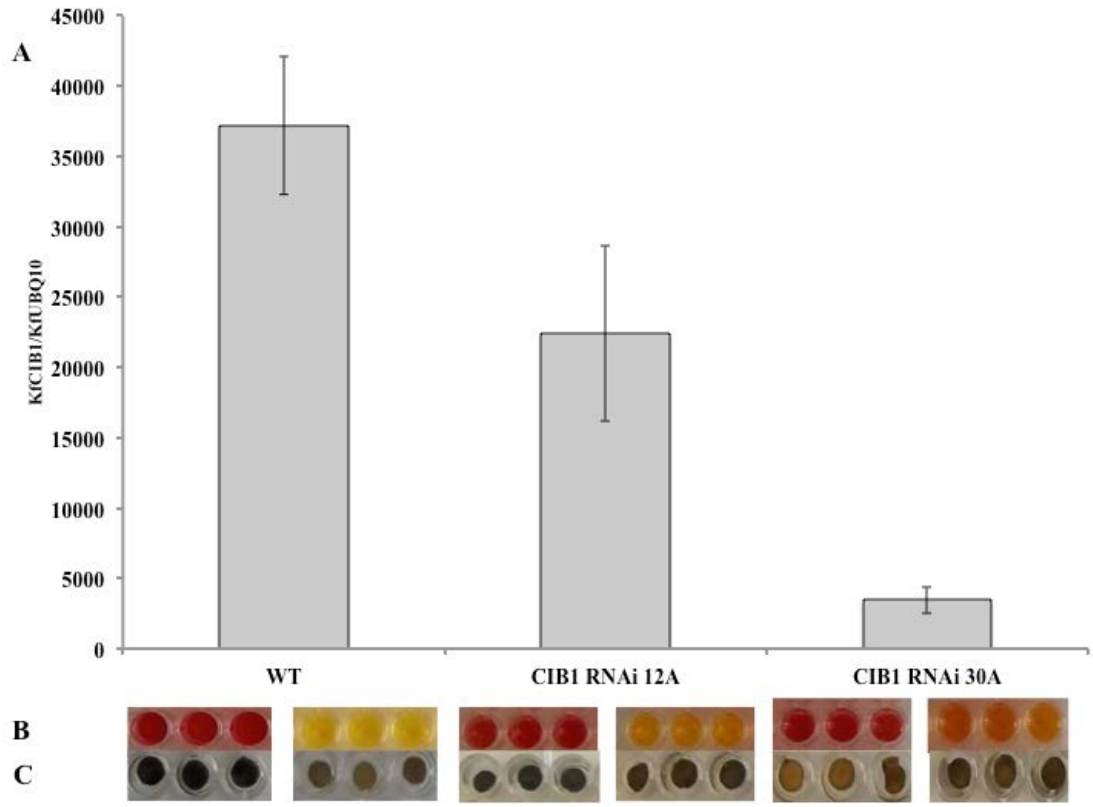

Figure 3.5. Transcript level, malate and starch results of KfCIB1 in Wild Type and the two chosen lines to be used for further analysis. **A.** Showing transcript level at the end of the dark period, when this transcript usually peaks. Shows varying level of transcript reduction in the chosen lines. **B.** Shows chlorophenol red pH indicator dye after leaf discs have been incubated and removed. Red indicates a pH of 6 or over, whereas yellow indicates pH4 or below, indicating high levels of malate are present. Light samples are on the left and dark samples are on the right. **C.** Levels of starch over the 24 hour cycle, showing light results to the left and dark results to the right.

Table 3.8. Summary of screening results for the two *KfCIB1_RNAi* knockdowns being taken forward for further analysis.

	CIB1 RNAi 12A	CIB RNAi 30A
Fold decrease of gene expression - Dark	1.7	10.8
Malate - Light	Low	Low
Malate - Dark	No Visible Difference	Low
Starch – Light	No Visible Difference	Low
Starch - Dark	High	No Visible Difference

3.2.2.3 KfCDF2

The sequence of *CYCLING DOF FACTOR2* (*KfCDF2*) from *K. fedtschenkoi* (KF17950 in the draft genome) has the highest hit with Arabidopsis ortholog At5g39660, also called *CDF2*. It belongs to the DOF family of plant transcription factors that possess a highly conserved DNA binding domain to enable control of gene expression. This family is involved in a wide variety of processes in *Arabidopsis*, such as defense responses, seed germination and photoperiodic flowering response (Chen *et al.,* 1996; Gualberti *et al.,* 2002; Corrales *et al.,* 2014). As this transcription factor is involved in many different processes, including flowering, then it is likely to be linked to the circadian clock in some manner, and as it is upregulated in CAM tissues by 6- to 11-fold, it may also signal to, or control some of the CAM biochemistry.

3.2.2.3.1 *KfCDF2* over-expresser lines

42 individual lines were created, of which 26 were shown to be over-expressers, with an average fold transcript increase of 284 at the end of the light period, and a 2.3 fold increase at the end of the dark. This overexpressed gene is another like *KfMYB439* over-expressing lines, which shows large variations in levels of over-expression, despite the CaMV35S promoter. An example of this is *KfCDF2_FL_13A,* which is the highest overexpressing line, where in the light the transcript level is raised by 443 fold, as wild type levels are very low at 10:00, and during the dark overexpresses by 2.4 fold. *KfCDF2_FL_69C* is the lowest overexpressing line, with only a 10.6 fold increase in transcript in the light (10:00) and a 1.95 fold increase during the dark (22:00; Table 3.9).

These lines seemed to show a range of phenotypes for sap pH and starch content in leaf discs (Table 3.9). 8 lines showed malate level decreases over light and dark periods, whilst only 2 lines showed consistent increases (Table 3.9). Again though, these are only crude assays, and will be repeated with a more precise enzyme assay on lines chosen to be taken forward for more in depth screening. Starch seemed slightly more consistent, where the majority had no large difference when compared to the wild type. Although, as mentioned before, this visual assay has difficulties at detecting high starch at the end of the light, because WT is already very dark at that point anyway, and it is also hard to detect low starch at the end of the dark, as wild type discs are already pale.

Table 3.9. Summary table for *KfCDF2_FL* screening results

Transgene being manipulated	Line	Transcript change relative to WT - Light	Transcript change relative to WT - Dark	pH relative to WT - Dusk	pH relative to WT - Dawn	Starch relative to WT - Dusk	Starch relative to WT - Dawn
KFCDF2 FL	1B	109	1.2	=	=	<	=
KFCDF2 FL	6A	146.5	1.9	<	<	=	=
KFCDF2 FL	13A	443	2.4	<	<	=	=
KFCDF2 FL	13B	2.7	1.5	=	=	=	<
KFCDF2 FL	14	415.7	1.82	<	=	<	<
KFCDF2 FL	14C	312	1.3	=	<	=	<
KFCDF2 FL	15B	65.3	1.73	=	<	<	<
KFCDF2 FL	15C	18.6	1.3	>	>	=	=
KFCDF2 FL	16C	10.39	2.44	<	<	=	=
KFCDF2 FL	17B	87.1	1.1	>	=	=	=
KFCDF2 FL	18B	154.7	1.7	>	<	=	>
KFCDF2 FL	21A	121.3	1.63	<	<	=	>
KFCDF2 FL	26C	93.2	1.1	<	<	<	<
KFCDF2 FL	28B	24.5	1.2	<	=	=	=
KFCDF2 FL	34A	19.4	1.7	<	=	=	<
KFCDF2 FL	34C	20.1	1.1	=	<	<	<
KFCDF2 FL	41A	42	1.1	<	<	<	=
KFCDF2 FL	41C	34.1	1.9	<	<	=	=
KFCDF2 FL	43A	356.4	5.3	<	=	=	=
KFCDF2 FL	43B	297.4	1.37	<	<	=	=
KFCDF2 FL	43C	346.5	5.4	=	=	=	<
KFCDF2 FL	53A	46.5	3.1	=	<	=	=
KFCDF2 FL	53B	57.3	1.62	>	>	<	<
KFCDF2 FL	59A	87.9	1.2	=	=	=	<
KFCDF2 FL	64C	432.9	1.0	<	=	=	=
KFCDF2 FL	69C	10.6	1.95	<	=	=	<

Results of fold reduction of transcript in relation to WT (10h light and dark), Malate (pH) and starch leaf disc screen at 11h Light and 11h Dark over a 12:12 LD cycles for the confirmed transgenic lines. All results are in comparison to an average wild type of three biological replicates. For pH, WT is less acidic at dusk and highly acidic at dawn, and so symbols represent lines showing higher, lower or equal acidity than wild type at that time point accompanied with either red indicating a decrease, yellow indicating no change and green showing an increase. For starch levels the symbols represent higher, lower or equal starch levels at dawn and dusk. At dawn starch is low in wild type and at dusk there are high levels present. < indicates a lower pH or a lower amount of starch present, > shows a higher value for the transgenic and = represents no difference to the WT.

Lines with dark grey shading represent lines that show the additional phenotype of photoperiod independent flowering. Light grey shading represents plants that did not

actually produce buds and flower, but their stem began to hook, showing the first stage of flowering initiation.

During the malate and starch screening process, plants that were not being entrained in the Snijder growth chambers were kept in the greenhouse which had supplementary lighting to ensure plants were kept in 16:8 L:D (long days) conditions to prevent flowering. Kalanchoe preferentially flower in short days (8:16 LD), but flowering can also be initiated in 12:12 LD conditions. Throughout the months of screening it became apparent that these lines flowered regardless of the photoperiod, as many confirmed overexpresser plants began to hook, bud and then flower in long days (Fig. 3.6). some lines did not get to actually produce buds and flowers, but still had hooked meristems, but did not seem to progress any further, and complete flowering (Tables 3.9 & 3.10). The stage reached in progression to flowering does appear to be dependent of overexpression level, as lines that produced flowers have a higher *KfCDF*2 transcript level at both time points (dark grey boxes, Table 3.9), compared to lines that only showed a hooked meristem (light grey boxes, Table 3.9). Furthermore, from the results for line *KfCDF2_FL_64C*, where there is no increase in *KfCDF2* expression during the dark, but a nearly 433-fold increase during the light, it suggests that increased expression in the light may be key for flowering induction regardless of photoperiod. No Wild Type lines showed any signs of flowering, which were grown along side the *KfCDF2_FL* lines. This was therefore another visual screen that could be used to easily identify overexpressing lines.

Figure 3.6. *KfCDF2* overexpressing plants in greenhouse conditions showing photoperiod insensitive flowering, surrounded by other lines, including wild type, which are showing only vegetative growth.

Table 3.10. Summary of *KfCDF2_F*L lines that have shown photoperiod insensitive flowering in normal non-flowering conditions (16:8 LD). Individual lines are highlighted in Table 3.9 in either light grey (hooked lines) or dark grey (flowering lines).

CDF2 FL	Successfully flowered	Hooked Meristem
Number of Lines	11/26	9/26
% of Lines	42	35
Fold Transcript Increase	L – 284 D – 2.3	L – 60 D – 1.7

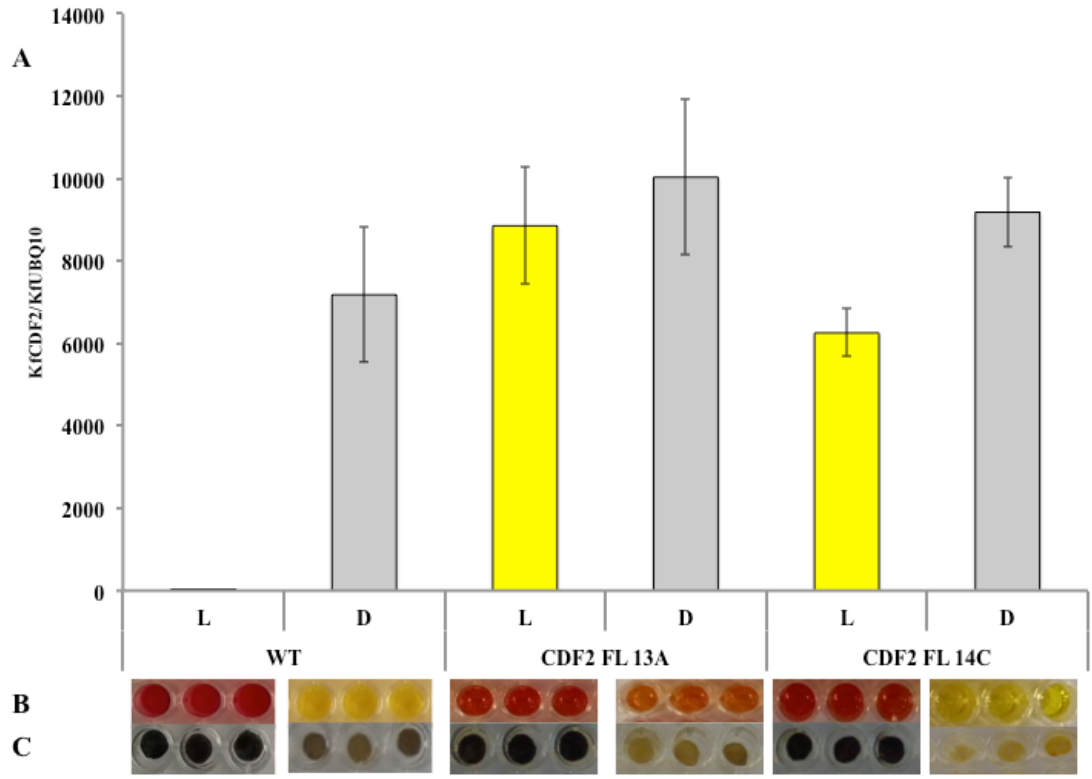

Figure 3.7. Transcript level, malate and starch results of KfCDF2 in Wild Type and the two chosen lines to be used for further analysis. **A.** Showing transcript level is naturally very low during the day, and peaks during the night. The overexpressing lines; 13C and 14C, show a drastic increase in light expression, and a smaller increase, during the dark. **B.** Shows chlorophenol red pH indicator dye after leaf discs have been incubated and removed. Showing accumulation and decarboxylation over light and dark. **C.** Levels of starch over the 24 hour cycle. At the end of light there is expected to be high starch due to photosynthesis (black disc), and at the end of the dark there should be low starch due to breakdown to produce PEP (Light coloured leaf discs).

The two lines chosen to be brought forward were *KfCDF2_FL_13A* and *14C*. Both lines were successful overexpressors of *KfCDF2*, and both of these lines had also successfully flowered in greenhouse conditions at different times. These plants also seemed interesting as some lines did not grow well and had smaller, thicker leaves, (*KfCDF2_FL_13A*) whereas others grew rapidly with large leaves (*KfCDF2_FL_14C*), and so one of each was chosen for further investigation (Fig. 3.7; Table 3.11).

Table 3.11. Summary of screening results for the two *KfCDF2_FL* lines overexpressers being taken forward for further analysis.

	CDF2 FL 13A	CDF2 FL 14C
Fold increase of gene expression - Light	442.7	312.7
Fold increase of gene expression - Dark	2.4	1.3
Malate - Light	Low	No Visible Difference
Malate - Dark	Low	Low
Starch – Light	No Visible Difference	No Visible Difference
Starch - Dark	Low	Low

3.2.2.3.2 *KfCDF2* RNAi lines

All 42 lines generated from this construct showed reduction in transcript level, giving this construct 100% transformation efficiency. The fold reductions did vary though from 1.01 to 3.1, but averaging at 1.6 less *KfCDF2* mRNA than the wild type. From the malate and starch screens, like with the *KfCDF2_FL* lines, there was little correlation between reduction in transcript level and changes seen in sap pH and starch levels (Table 3.12). The trend from *KfCDF2* reduction is, a decrease in malate at the end of the light (22 lines) and dark (19 lines; Table 3.12). The majority of the lines appear to have similar starch levels to the wild type, with 30 lines showing no difference to wild type at then end of the light period, and 23 lines with no difference at the end of the dark, suggesting there are not many lines where starch metabolism has been greatly disrupted, or just that a more accurate and sensitive test is needed.

Table 3.12. Summary table for *KfCDF2_RNAi* screening results

Transgene being manipulated	Line	Transcript change relative to WT - Dark	pH relative to WT - Dusk	pH relative to WT - Dawn	Starch relative to WT - Dusk	Starch relative to WT - Dawn
KfCDF2 RNAi	1A	1.5	=	=	<	<
KfCDF2 RNAi	1B	1.7	=	<	<	=
KfCDF2 RNAi	1C	1.9	<	>	=	=
KfCDF2 RNAi	2A	1.6	=	<	=	=
KfCDF2 RNAi	2B	1.7	=	<	=	=
KfCDF2 RNAi	2C	1.7	=	<	=	=
KfCDF2 RNAi	7B	1.4	=	=	=	=
KfCDF2 RNAi	7C	2.5	=	<	=	=
KfCDF2 RNAi	8A	1.7	<	<	<	=
KfCDF2 RNAi	9B	2.7	>	=	=	=
KfCDF2 RNAi	18C	1.8	<	=	=	=
KfCDF2 RNAi	19A	1.4	<	<	=	<
KfCDF2 RNAi	20A	1.3	>	=	=	=
KfCDF2 RNAi	22A	1.2	<	<	=	<
KfCDF2 RNAi	24A	1.1	=	<	<	<
KfCDF2 RNAi	24B	1.5	=	<	<	<
KfCDF2 RNAi	26B	1.8	<	<	=	<
KfCDF2 RNAi	26C	1.7	<	<	=	<
KfCDF2 RNAi	28A	1.9	<	=	=	<
KfCDF2 RNAi	28B	1.9	<	<	<	<
KfCDF2 RNAi	36A	1.8	<	=	=	=
KfCDF2 RNAi	36C	1.4	<	=	=	<
KfCDF2 RNAi	38B	1.9	<	>	=	=
KfCDF2 RNAi	41B	1.6	=	=	=	<
KfCDF2 RNAi	41C	3.1	<	=	=	<
KfCDF2 RNAi	42A	1.8	=	<	=	<
KfCDF2 RNAi	42B	1.6	>	>	<	<
KfCDF2 RNAi	42C	1.7	>	=	<	<
KfCDF2 RNAi	49A	1.2	<	=	<	=
KfCDF2 RNAi	49B	1.3	<	=	=	=
KfCDF2 RNAi	49C	1.3	<	=	=	=
KfCDF2 RNAi	50A	1.5	<	=	=	<
KfCDF2 RNAi	50B	1.7	>	=	=	=
KfCDF2 RNAi	52B	1.7	=	<	<	<
KfCDF2 RNAi	52C	1.1	=	<	<	<
KfCDF2 RNAi	55B	2.1	<	<	=	=
KfCDF2 RNAi	55C	1.6	=	=	=	<
KfCDF2 RNAi	56A	1.2	<	<	=	=
KfCDF2 RNAi	56B	1.2	<	<	=	=

KfCDF2 RNAi	56C	1.5	<	=	=	=
KfCDF2 RNAi	57A	1.2	<	=	=	<
KfCDF2 RNAi	57B	1.8	>	=	<	=

Results of fold reduction of transcript in relation to WT (10h light and dark), Malate (pH) and starch leaf disc screen at 11h Light and 11h Dark over a 12:12 LD cycles for the confirmed transgenic lines.All results are in comparison to an average wild type of three biological replicates. For pH, WT is less acidic at dusk and highly acidic at dawn, and so symbols represent lines showing higher, lower or equal acidity than wild type at that time point accompanied with either red indicating a decrease, yellow indicating no change and green showing an increase. For starch levels the symbols represent higher, lower or equal starch levels at dawn and dusk. At dawn starch is low in wild type and at dusk there are high levels present. < indicates a lower pH or a lower amount of starch present, > shows a higher value for the transgenic and = represents no difference to the WT.

It is also of note, that none of these lines flowered, despite being in the same growth conditions as the *CDF2_FL* lines.

The lines *KfCDF2_RNAi_19A* and *KfCDF2_RNAi_26B* were chosen to be taken forward for further analysis (Fig. 3.8; Table 3.13). These lines were chosen because *KfCDF2_RNAi_26B* was one of the lines with the highest knockdowns, which showed differences in both malate and starch metabolite levels. *KfCDF2_RNAi_19A* was also chosen, because it showed a more intermediate level of transcript reduction, yet still showed the same results as *KfCDF2_RNAi_26B* for pH and starch leaf screens. This would therefore be interesting to determine if dosage effects will play a role in severity of perturbations of these metabolite pools.

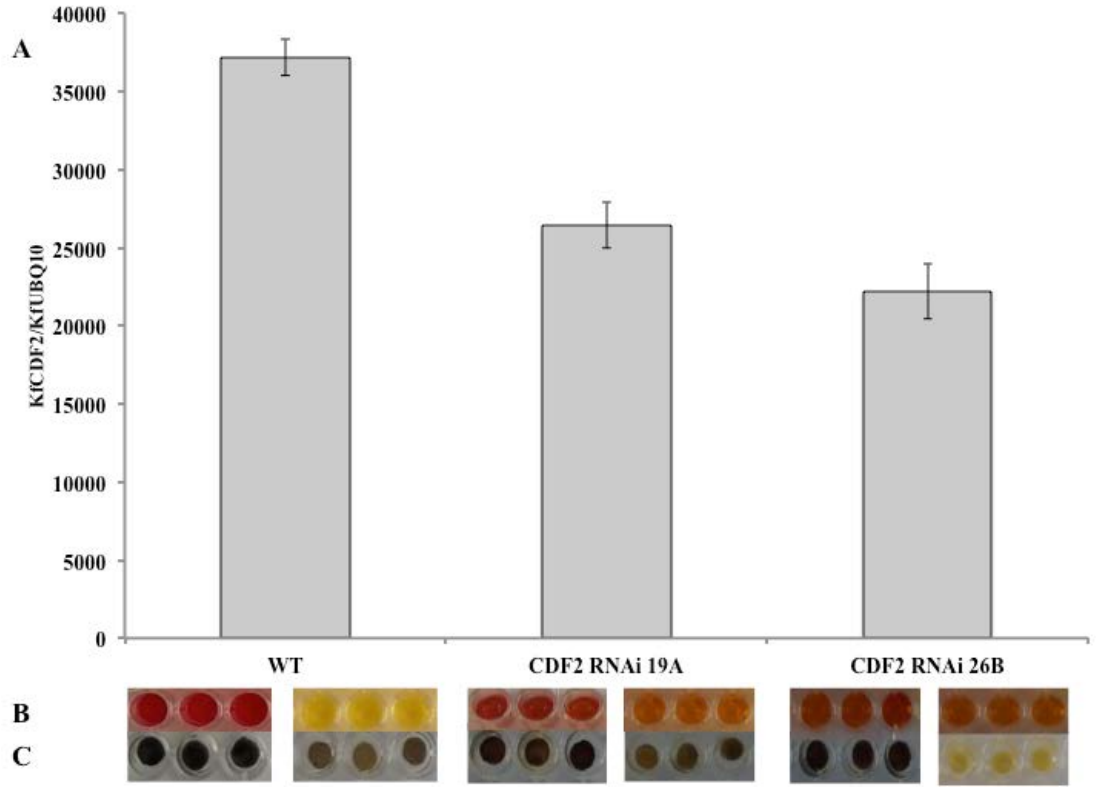

Figure 3.8. Transcript level, malate and starch results of *KfCDF2_RNAi* in Wild Type and the two chosen lines to be used for further analysis. **A.** Showing transcript level at the end of the dark, where the highest level is seen in WT. **B.** Shows chlorophenol red pH indicator dye after leaf discs have been incubated and removed. End of light samples are on the left, and end of dark samples are on the right. **C.** Iodine stain to estimate starch levels. At the end of light there is much starch produced in Wild Type from photosynthesis and so leaves appear black in the iodine stain (left). At the end of the dark, starch has been used for PEP synthesis, and so lighter coloured discs should be seen (right). Differences therefore suggest improper starch cycling.

Table 3.13. Summary of screening results for the two *CDF2_RNAi* knockdowns being taken forward for further analysis.

	CDF2 RNAi 19A	CDF2 RNAi 26B
Fold decrease of gene expression - Dark	1.4	1.8
Malate - Light	Low	Low
Malate - Dark	Low	Low
Starch – Light	No Visible Difference	No Visible Difference
Starch - Dark	No Visible Difference	Low

3.3 Discussion

The over-arching aim set at the start of this PhD was to define genes involved in the circadian control and optimisation of CAM and to characterise their functions. This chapter summarises the screening of a large population of primary transformants of *K. fedtschenkoi* in which the candidate CAM-induced, clock-controlled TFs, *KfMYB439, KfCIB1* and *KfCDF2* were either down-regulated with RNAi constructs, or up-regulated constitutively with 35S promoter-driven constructs. These TF genes were identified originally using SOLiD RNA-Seq sequencing, and transformed *K. fedtschenkoi* lines were generated by Jana Knerova as part of a previous BBSRC-funded functional genomics of CAM and its circadian clock control in *K. fedtschenkoi* project. To identified the best lines for further study in this PhD project, initial phenotypic and transcript level testing was conducted on over 200 primary transformants. Semi-quantitative RT-PCR was used to estimate the level of the transcript targeted by the transgene, and malate (pH) and starch assays were conducted to look for major perturbations in CAM biochemistry. These pre-screens revealed that the overexpression or knockdown of these TF genes did affect the size of key metabolite pools associated with CAM in at least a subset of the available transgenic lines. These pre-screens thus formed a sound basis for selecting a relatively small number of lines for each TF to be taken forward for more detailed characterization of phenotypes associated with CAM and its control by the circadian clock (Table 3.14).

As reciprocal cycling of the metabolic pools of malate and starch is a key phenotype of CAM, the high throughput nature of these malate and starch leaf disc staining assays was very desirable. It had been demonstrated previously that these staining tests were able to detect large and consistent differences in CAM mutants of *M.*

crystallinum, and that they were also a powerful pre-screen for transgenic lines of *K. fedtschenkoi* in which the key CAM metabolic enzymes *KfPPDK* and *KfNAD-ME1* were down-regulated (Cushman *et al.,* 2008a; Dever *et al.,* 2015).

However, with the *KfMYB439*, *KfCIB1* and *KfCDF2* RNAi and over-expressor lines studied here, detection of variations in dawn and dusk malate and starch levels using these leaf disc stains led to more subtle and less consistent results. This may be due to the targeted genes not playing such a major role in CAM. Based on the findings presented here, it seems that these screens are only really appropriate when the targeted gene is a major CAM component. Based on this limitation, the selection of the lines for further, more detailed phenotypic characterisation throughout the remainder of this thesis relied most heavily on the fold-change in transcript level between the transgenic line and the wild type. However, lines that also showed large changes in the pH and starch staining results were also kept for further analysis where this correlated well with changes in the targeted transcript level (Table 3.14).

Table 3.14. Summary table of lines chosen to be taken forward for further phenotypic analysis.

Line	Approximate fold gene expression change - Light	Approximate fold gene expression change - Dark	Malate level - Light	Malate level - Dark	Starch level - Light	Starch level - Dark
MYB439 FL 37C	+ 38.0	+ 1.1	Low	Low	Low	Low
MYB439 FL 54B	+ 24.4	+ 1.7	High	High	Low	Low
MYB439 RNAi 3C	-	- 1.5	High	Low	Low	No Change
MYB439 RNAi 10B	-	- 1.3	High	High	Low	Low
CIB FL 5C	+ 6.4	+ 1.6	Low	Low	Low	No Change
CIB FL 11B	+ 3.1	+ 1.3	High	High	No Change	No Change
CIB RNAi 12A	-	- 1.7	Low	No Change	No Change	High
CIB RNAi 30A	-	- 10.8	Low	Low	Low	No Change
CDF2 FL 13A	+ 442.7	+ 2.4	Low	Low	No Change	Low
CDF2 FL 14C	+ 312.7	+ 1.3	No Change	Low	No Change	Low
CDF2 RNAi 19A	-	- 1.4	Low	Low	No Change	No Change
CDF2 RNAi 26B	-	- 1.8	Low	Low	No Change	Low

A key priority for the subsequent phenotypic characterization of these lines was to determine the levels of malate and starch at dawn and dusk more precisely by using enzyme-linked spectrophotometric assays that are quantitative. This allowed the testing of the preliminary results from the pH and starch staining tests performed on the leaf discs. The results of the quantitative malate and starch assays are reported in the following chapters, which describe the detailed phenotypic characterization of each population of selected transgenic lines.

Furthermore, the *KfCDF2_FL* lines were found to display photoperiod-insensitive flowering in long days under 16:8 greenhouse conditions. This supports the

identification of this gene as a component of a circadian clock output pathway. In Arabidopsis, the *CDF* family of TFs, including *CDF2*, have been shown to be involved in a wide variety of biological processes including carbon metabolism (Yanagisawa *et al.,* 1998), phytochrome signaling (Park *et al.,* 2003), as well as flowering responses (Imaizumi *et al.,* 2005; Fornara *et al.,* 2009) all of which are circadian clock linked. Thus it is possible that this gene may have been recruited to a role in the circadian regulation of CAM, as suggested originally by the SOLiD RNA-Seq data, which showed *KfCDF2* to be strongly induced in CAM leaves relative to C_3 leaves.

Finally, papers such as Farinas & Mas, (2011) and Hsu *et al.,* (2013), provide compelling evidence that the RVE family of TFs, especially RVE 4, 6 and 8 (KfMYB439s ortholog in *Arabidospsis*), actually play roles in the core circadian clock, functioning as evening gene activators, and playing a antagonistic role with CCA1/LHY. *CCA1/LHY* are also part of the RVE family, suggesting the possibility that this gene is interacting with evening-phased CAM genes, and may indeed be involved in the circadian clock.

Chapter 4 – Characterisation of the CAM-associated phenotypic changes in transgenic RNAi and over-expression lines of *K. fedtschenkoi* with altered expression levels of the CAM-induced and clock-controlled transcription factor *MYB-LIKE439 (MYB439)*

4.1 Introduction

KfMYB439 (KF79015 in draft genome assembly) was chosen because, as explained in Chapter 3, it was found to be under light/dark and circadian clock control, and also showed strong induction in CAM leaves.

In terms of orthology, *KfMYB439*, had greatest percentage identity with the Arabidopsis gene *HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN/ REVEILLE 8* (*HSP, RVE8*, gene ID: At3g09600), which is involved in regulating the circadian clock by modulating histone 3 (H3) acetylation (Farinas *et al.,* 2011). *KfMYB439* also had a high percentage pairwise identity to a second gene called LHY/CCA1 LIKE 1, REVEILLE 4 (LHY/CCA1-LIKE 1, RVE4, gene ID: AT5g02840). Thus, in addition to being a CAM-induced and circadian-clock controlled gene, *KfMYB439* was also shown to be similar in sequence to key known circadian clock genes belonging to the *REVEILLE* family in Arabidopsis. These facts support the proposal that this transcription factor is likely to play a role in oscillations in the circadian clock in *K. fedtschenkoi*, in addition to any role it may play in the output pathway that links the central oscillator to the circadian control of CAM.

4.1.1 MYB Transcription Factors

The MYB family of transcription factors is found in all eukaryotes, and is large and functionally diverse. Plants are the kingdom that have the highest diversity of this family, likely due to selective expansion, with between 100 and 200 MYB family members commonly found in individual plant species (Dubos *et al.,* 2010). In animals, the majority of MYB genes have three MYB repeats, whereas in plants, the

majority tend to have two. Interestingly, whilst animal MYB proteins have diverged from a single common ancestor, plant MYBs appear to be a polyphyletic group, related only within the DNA-binding domain (Rosinski *et al.,* 1998). Related binding domains enable these transcription factors to bind similar target motifs in DNA, and therefore carry out similar regulatory roles (Table 4.1).

Possibly one of the most well-known MYB-like transcription factors is the central circadian clock gene *CCA1*. CCA1 is essential to promoting deacetylation of Histone 3 (H3), which prevents *TOC1* mRNA accumulation. It has also been shown to bind to the evening element (EE; AAAATATCT) of evening phased genes like *TOC1* to repress expression (Farinas *et al.,* 2011). It has also been shown that overexpression of CCA1, leads to decreases in H3 acetylation and *TOC1* mRNA (Farinas *et al.,* 2011).

Table 4.1. Table of the four groups of plant MYB transcription factors and their main functions.

MYB Family Group	Main Functions
4R-MYB	Unknown. Only one has been found.
R1R2R3-type MYB	Cell cycle control.
R2R3-type MYB	Flavonoid and Anthocyanin biosynthesis, pollen wall development, cellulose and lignin biosynthesis, biotic and abiotic stress responses.
Single/Partial MYB - 'MYB Related'. (Many subclasses)	Cellular morphogenesis, secondary metabolism, central circadian clock control, Chloroplast development and abiotic stress.

4.1.2 MYB Structure

Each MYB domain (R; Fig. 4.1) consists of up to 4 varying amino acid sequence repeats (of around 50 amino acids), which form three α-helicies with a protein structure of helix-helix-turn-helix (Martin *et al.,* 1997).

There are 4 plant MYB protein factor classes, which are classified depending on the number of repeats, and the type of domain (R; Table 4.1). The domain names R1, R2 and R3 come from the Human oncogene c-MYB, and all MYB domains are classified due to their similarity to R1-3 (Fig. 4.1).

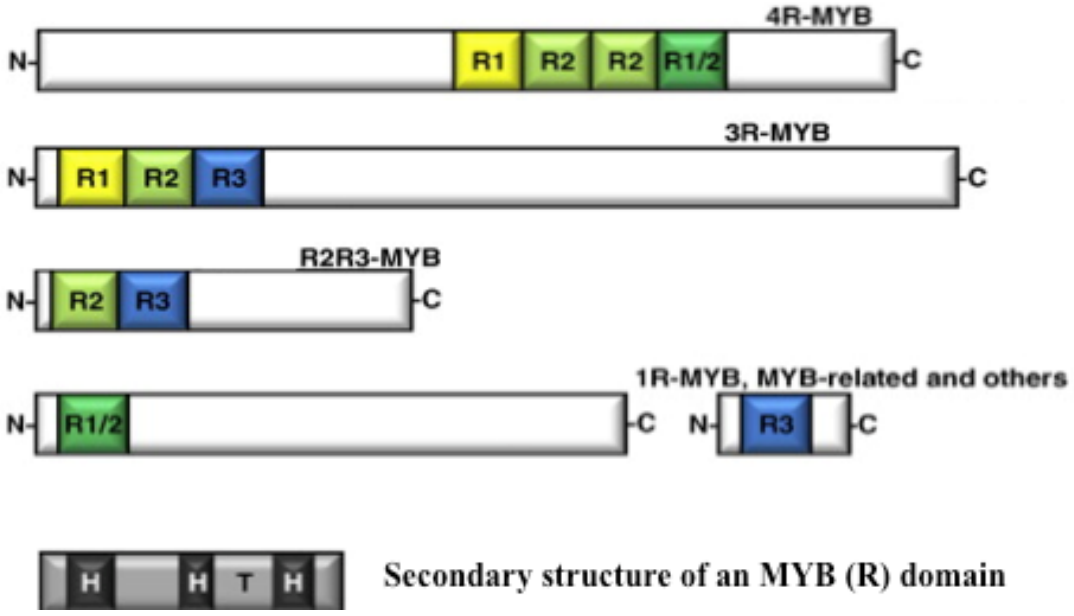

Figure 4.1 The 4 groups of MYB transcription factors in plants, and also their secondary structure; helix, helix-turn-helix. Reproduced from Dubos *et al*., (2010).

4.1.3 Single MYB-repeat proteins

Single MYB-repeat proteins predominantly bind to telomeric sequences or histones, many function in the circadian clock, and all display similar sequence identity (Martin *et al.,* 1997).

In rice, it has been shown that single MYB-repeat factors are able to form dimers (Lu *et al.,* 2002). The composition of these dimers determines different binding affinities,

and therefore different levels of activation for downstream gene expression. Furthermore, in *Arabidopsis*, single MYB-repeat proteins have been shown to bind within areas that 2-repeat MYBs (R2R3-MYBs) bind, and so function as competitors for the same DNA targets (Feldbrugge *et al.,* 1997; Liao *et al.,* 2008). This provides a way for these single MYB-repeat transcription factors, and their range of combinatorial dimers, to fine-tune gene expression through their different binding affinities and activation abilities, and also by competition for binding. These factors vary according to different levels of these MYBs in a cell (Feldbrugge *et al.,* 1997; Liao *et al.,* 2008).

Single-MYB repeat proteins have been shown to bind DNA in a different manner to other MYBs. They have been shown to form homo- and heteromultimers, whereas the multiple MYB domains in the rest of the MYB transcription factor family enable them to function alone (Hwang *et al.,* 2001).

4.1.4 *KfMYB439*

From comparison of the sequences using the Geneious alignment option (with Blosum62 as the cost matrix); a Global Multiple sequence alignment of the amino acid sequences for *A. thaliana HSP*, *LHY* and *CCA1*, with those of *LHY, CCA1* and *MYB439* from *K. fedtschenkoi*, confirmed the close relationship of *KfMYB439* to these known clock genes possessing a single MYB-repeat (Fig 4.2). This analysis confirmed that all of these genes belong to the REVEILLE (RVE) sub-family of Arabidopsis MYBs. They all possessed the conserved SHAQKY amino acid sequence in the DNA binding domain, and also the conserved residues directly upstream of it (Fig 4.2). This family is composed of 11 proteins in *Arabidopsis*, including RVE8

(At3g09600), which was the best BLAST hit when *KfMYB439* was used as a query against the Arabidopsis genome using the TAIR website (www.arabidopsis.org).

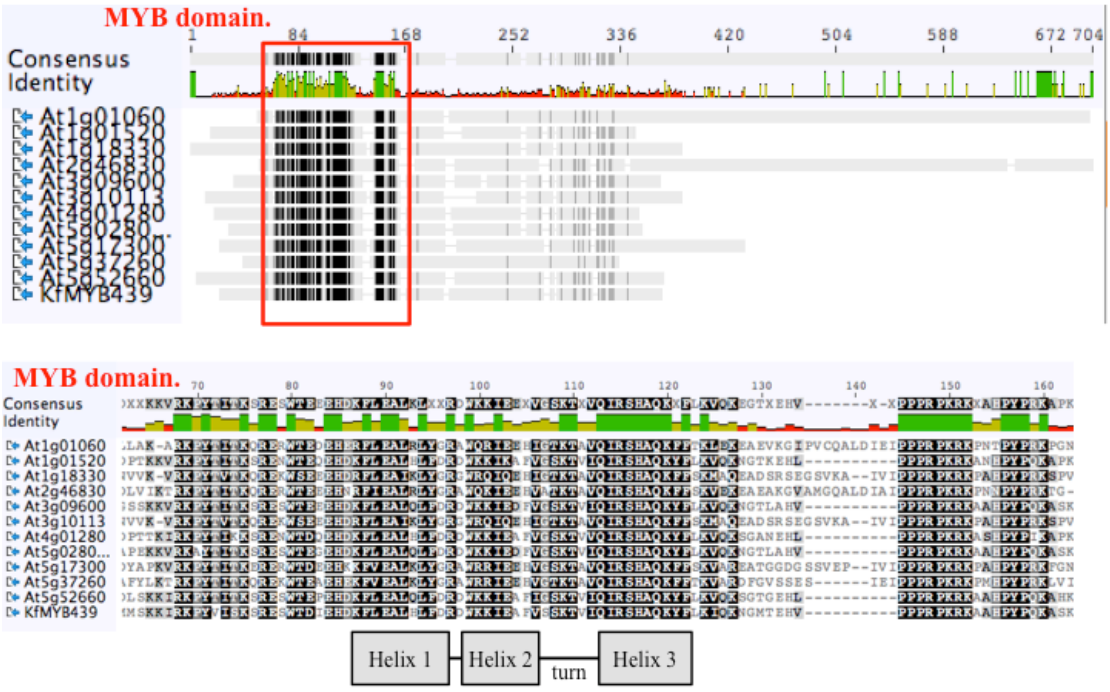

Figure 4.2. Global multiple sequence alignment (using Geneious alignment using Blosum62) showing the conservation of the single MYB-domain in KfMYB439 relative to the corresponding domain in the *REVEILLE* family from Arabidopsis. Information on the conserved areas in SHAQKY MYB transcription factors was obtained from Rose *et al*., 1999. LHY = AT1G01060, CCA1 = AT2G46830, RVE1 = AT5G17300, RVE2 = AT5G37260, RVE3 = AT1G01520, RVE4 = AT5G02840, RVE5 = AT4G01280, RVE6 = At5G52660, RVE7 = AT1G18330, RVE8 = AT3G09600 and MYB-LIKE HOMEODOMAIN SUPERFAMILY PROTEIN = AT3G10113.

The third helix is referred to as the 'recognition helix', as this is the helix responsible for binding DNA in the major groove, and intercalating there. *KfMYB439* has a very similar 'recognition helix' to central clock genes *CCA1* and *LHY*, with the exception of one amino acid within the SHAQK[F/Y]F helix, suggesting potential competition at binding sites, and also overlap of function, as shown with other RVE family members: CCA1/LHY and RVE2 (Zhang *et al.,* 2007). This highlights a key challenge with the characterisation of the function of single RVE-family TF though,

as due to the DNA binding sites being similar, proteins can potentially compete with different MYBs. This has been shown previously in *Antirrhinum,* where two flower-specific MYBs, MYB305 and MYB340, are involved in flavonoid metabolism (Moyano *et al.,* 1996). MYB340 is a stronger activator of other genes required in the flavonoid pathway, but MYB305 protein is more abundant, resulting in more MYB305 being able to bind, and so gating the signal level. Levels can be modulated though depending on flowering status of the plant (Moyano *et al.,* 1996). This therefore suggests that net transcription activation may depend on the relative amounts of the different proteins and also their different abilities to activate transcription.

4.1.5 REVEILLE family of MYB Transcription Factors.

A large volume of research has been conducted on the *RVE* family of transcription factors in *Arabidopsis*. The discovery of the founding members of this 11 member MYB sub-family in *Arabidopsis*, namely CCA1 and LHY, as key components of the core circadian clock mechanism fuelled intense interest in understanding the functions of the other 9 RVE genes (Perales *et al.,* 2007; Zhang *et al.,* 2007; Rawat *et al.,* 2009; Farinas *et al.,* 2011; Hsu *et al.,* 2013). RVE2 and RVE7 have been shown to be part of the regulatory feedback loops of the clock, and are closely linked to the core circadian oscillator (Zhang *et al.,* 2007). RVE1 has been found to not be essential to the clock, and works more in output pathways, such as modulating growth by controlling auxin levels during the day (Rawat *et al.,* 2009). CCA1 and its closely related homolog LHY both regulate both morning and evening expressed genes due to being core clock components (Harmer 2009). CCA1 has been found to be key in regulating of the acetylation of Histone 3 (H3) at the *TOC1* promoter, and therefore

modulating *TOC1* expression (Farinas *et al.,* 2011). Specifically, CCA1 has been shown to promote deacetylation, and thereby inhibits *TOC1* transcription (Perales *et al.,* 2007). RVE8 on the other hand has been shown to play an antagonistic role to CCA1, and has been shown to be key for acetylation of the *TOC1* promoter to occur, resulting in its repression (Farinas *et al.,* 2011). These antagonistic roles within the RVE family are thought to be the reason why *ccal/lhy* double mutants are not completely arrhythmic (Hsu *et al.,* 2013). It has also been demonstrated that CCA1 and RVE8 bind to the EE in the *TOC1* promoter and other evening phased genes, which leads to DNA binding competition between these RVE family members (Farinas *et al.,* 2011).

Also, amongst the 11 RVE members in Arabidopsis, it has been shown that there is functional redundancy between members. This redundancy is well defined between CCA1 and LHY (Mizoguchi *et al.,* 2002), but has also been demonstrated between RVE4, RVE6 and RVE8. In *Arabidopsis*, only the *rve4, 6, 8* triple mutant showed a severe circadian phenotype (Hsu *et al.,* 2013).

Interestingly, out of the 11 RVE family members, 5 possess a second conserved domain outside of the MYB domain called the LHY/CCA1-LIKE (LCL) domain, which is shown in Figure 4.2 (Schmied *et al.,* 2005). RVEs 1, 2, and 7, plus CCA1 and LHY do not have this domain, but RVE8 does (LCL-5). This may explain why RVE8 binds to similar promoter motifs, due to its MYB domain, but works antagonistically to CCA1, perhaps due to the LCL domain.

4.2 Results.

4.2.1 Bioinformatic characterisation of *KfMYB439*.

4.2.1.1. Identification of conserved domains within the MYB TF family.

To try and better understand the potential function of *KfMYB439* in gene expression regulation, the coding sequence was used as the query for a BLAST search against the TAIR10 Transcripts (-introns, +UTRs) (DNA) sequences (Table 4.2; www.arabidopsis.org). All of the top BLAST hits confirmed *KfMYB439* as a member of the RVE family, as was already suggested after annotation of the initial SOLiD sequencing data, and confirmed in Figure 4.2. The RVE family is composed of 11 proteins in Arabidopsis, and of these, *RVE8,* a member of the LCL sub-family, was the highest scoring BLAST hit for *KfMYB439* (Table 4.2). As discussed in the introduction, the LCL subfamily appear to have opposing functions to other members of the RVE family, *CCA1/LHY*, and so it was important to confirm in more detail which members of the RVE family *KfMYB439* was most closely related to.

Table 4.2. BLASTN results produced using NCBI BLAST2.2.8 (https://www.arabidopsis.org/Blast/index.jsp), showing highest hits produced. All high scoring hits can be seen to be members of the MYB-REVEILLE transcription factor family.

At accession number	Gene name	Symbol	Bit score	E value
AT3G09600.2	HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN	RVE8/LCL5	60	5e-08
AT3G09600.1	HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN	RVE8/LCL5	60	5e-08
AT5G02840.3	LHY/CCA1-LIKE 1	RVE4/LCL1	48	2e-04
AT5G52660.2	HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN	RVE6/LCL2	48	2e-04
AT5G02840.2	LHY/CCA1-LIKE 1	RVE4/LCL1	48	2e-04
AT5G02840.1	LHY/CCA1-LIKE 1	RVE4/LCL1	48	2e-04
AT5G52660.1	HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN	RVE6/LCL2	48	2e-04
AT4G01280.2	HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN	RVE5/LCL4	48	2e-04
AT4G01280.1	HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN	RVE5/LCL4	48	2e-04
AT1G01520.1	HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN	RVE3/LCL3	42	0.013
AT1G01060.1	LONG ELONGATED HYPOCOTYL	LHY	40	0.050

Many published reports have argued that plant single MYB-repeat TFs show little homology outside the MYB domain. However, Schmied and Merkle showed that a subset of plant single MYB-repeat TFs possess a conserved motif beyond the MYB-repeat (2005). Five of the 11 genes in the Arabidopsis RVE family shared a second domain with high homology, the LCL domain, which central clock genes *LHY* and *CCA1* did not possess. Sequences were obtained from TAIR for the 5 RVEs in the LCL group, namely At1g01520 (RVE3), At3g09600 (RVE8), At4g01280 (RVE5), At5g02840 (RVE4) and At5g52660 (RVE6). To determine whether or not *KfMYB439*

possessed the LCL-specific sub-domain, full-length amino acid sequences were aligned using the default Geneious Pro 4.5.5 global alignment tool; Blosum62 (Fig. 4.3).

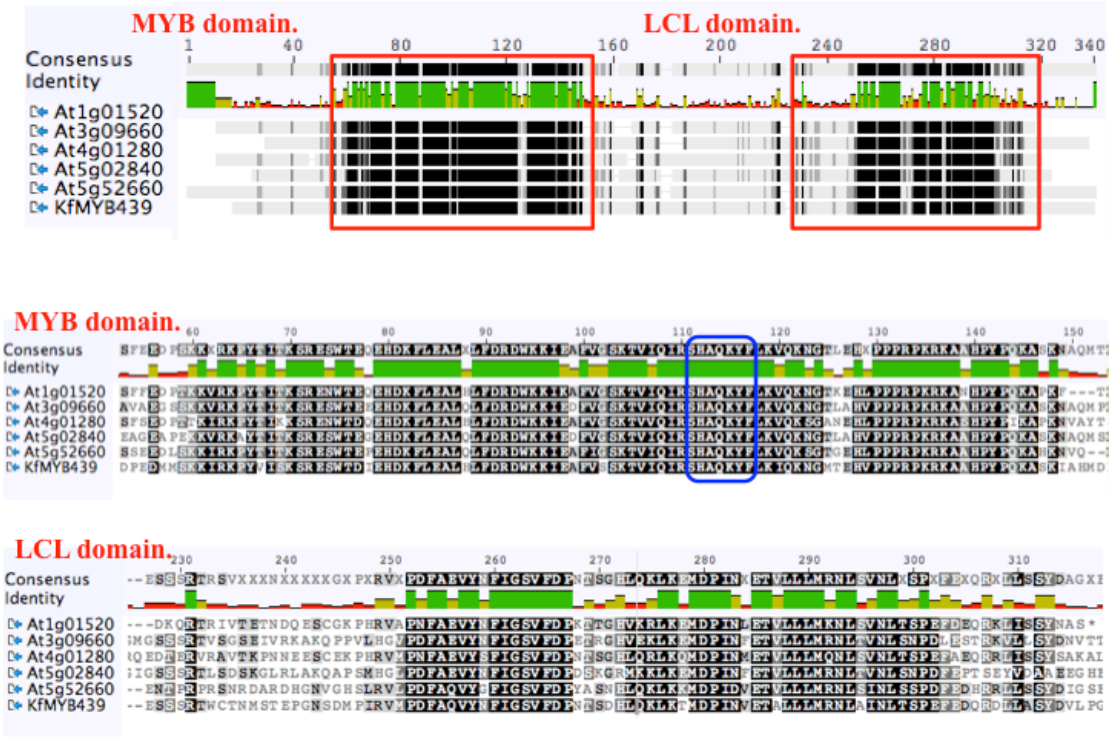

Figure 4.3. Geneious global alignments showing that *KfMYB439* contains both a conserved MYB domain and also a conserved LCL domain, both within the red boxes (upper panel). The middle and lower panels show zoomed in views for the MYB and LCL domain respectively. At1g01520 (RVE3), At3g09600 (RVE8), At4g01280 (RVE5), At5g02840 (RVE4) and At5g52660 (RVE6). The SHAQKYF amino acid sequence, which is possessed by all RVE TFs is highlighted within the blue box in the MYB domain.

Figure 4.3 shows that both the MYB and LCL conserved domains are present in *KfMYB439*, placing *KfMYB439* in the LCL subfamily, and also supporting the proposal that it could perform very similar functions to its top hit in *Arabidopsis*, *RVE8*.

4.2.1.2. Regulation of the *K. fedtschenkoi* RVE family in SOLiD RNA-seq quantitative transcriptome analysis

From the SOLiD RNA-seq dataset cataloguing the abundance of transcripts in C_3 and CAM leaves of *K. fedtschenkoi* sampled at 4 h intervals over a 12:12 light/ dark cycle, all 9 RVE family members identified in the *K. fedtschenkoi* genome were detected as being expressed. Two genes were identified as orthologs of CCA1/ LHY (KF02345 and KF106790), four had greatest homology to *Arabidopsis RVE1* (KF37745, KF84360, KF57920 and KF73515), and three were orthologs of Arabidopsis *RVE8* (KF29650, KF130900 and KF79015/ KfMYB439). All of the *K. fedtschenkoi* RVE family were expressed in both C_3 (LP1) and CAM (LP6) tissues, with an increase in expression level in CAM tissues (Fig. 4.4, Table 4.3). Three genes displayed greater fold-induction in LP6 relative to the rest of the family, suggesting these could potentially be of greater importance in CAM development and regulation, namely KF84360 (RVE1-like), KF37745 (RVE8-like) and KF79015 (RVE8-like/KfMYB439) (Table 4.3). KF130900 (pink line, Fig 4.4) appears to change very little throughout the light dark cycle, which is interesting, as all other lines are strongly circadian regulated. This gene also shows the lowest fold increase from C_3 to CAM leaves with only a value of 1.19 (Table 4.3).

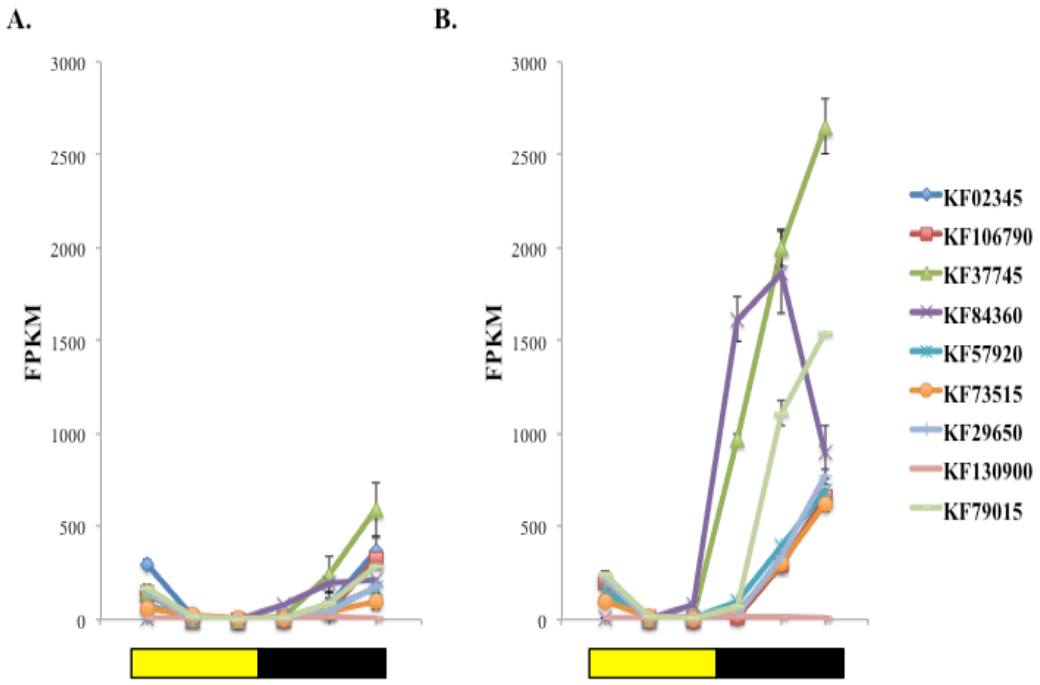

Figure 4.4. C_3 to CAM induction and light/ dark regulation of RVE-like genes identified in the SOLiD RNA-seq dataset for C_3 and CAM leaves of *K. fedtschenkoi* sampled at 4 h intervals over the light/ dark cycle. Transcript abundance of each gene is shown (FPKM). KF02345 and KF106790 represent CCA1/LHY, KF37745, KF84360, KF57920, KF73515 all represent RVE1-like genes, and KF29650, KF130900 and KF79015 all represent RVE8-like genes, with KF79015 being the gene code of *KfMYB439*. **A.** C_3 (LP1) *RVE* transcript abundance rhythms, showing low and similar expression levels for all RVEs. **B.** CAM (LP6) *RVE* transcript abundance rhythms over the light/dark cycle. Whilst all *RVE*s increased in transcript abundance in LP6 coincident with CAM, some RVEs showed a much higher induction in CAM tissues, namely RVE1-like genes KF37745 and KF84360, and also KF79015 (KfMYB439/RVE8-like).

Table 4.3. Regulation of the *K. fedtschenkoi* RVE family in C_3 and CAM leaves.

Gene Code	Gene Name	Fold Increase for C3 to CAM - Light	Fold Increase for C3 to CAM - Dark	Peak Expression Time	Fold increase of C3 to CAM - Peak expression time
KF02345	CCA1_LHY	0.92	4.49	22:00	1.82
KF106790	CCA1_LHY	1.37	10.41	22:00	2.01
KF37745	RVE1-like	2.63	22.32	22:00	4.48
KF84360	RVE1-like	10.12	11.20	18:00	9.67
KF57920	RVE1-like	5.21	11.09	22:00	4.01
KF73515	RVE1-like	0.87	6.28	22:00	6.39
KF29650	RVE8-like	1.10	10.41	22:00	4.28
KF130900	RVE8-like	1.31	1.37	22:00	1.19
KF79015	RVE8-like / KfMYB439	1.27	10.62	22:00	5.33

Fold-induction in CAM leaves (LP6) relative to C_3 leaves (LP1) was calculated for each RVE gene. Eight of the *RVE*s peaked at the same time,22:00 (2 h before dawn) but KF84360 (RVE1-like) peaked at 18:00, 4 h earlier, in the middle of the dark period.

4.2.1.3. Impact on the RVE family, when other genes are manipulated in *K. fedtschenkoi*

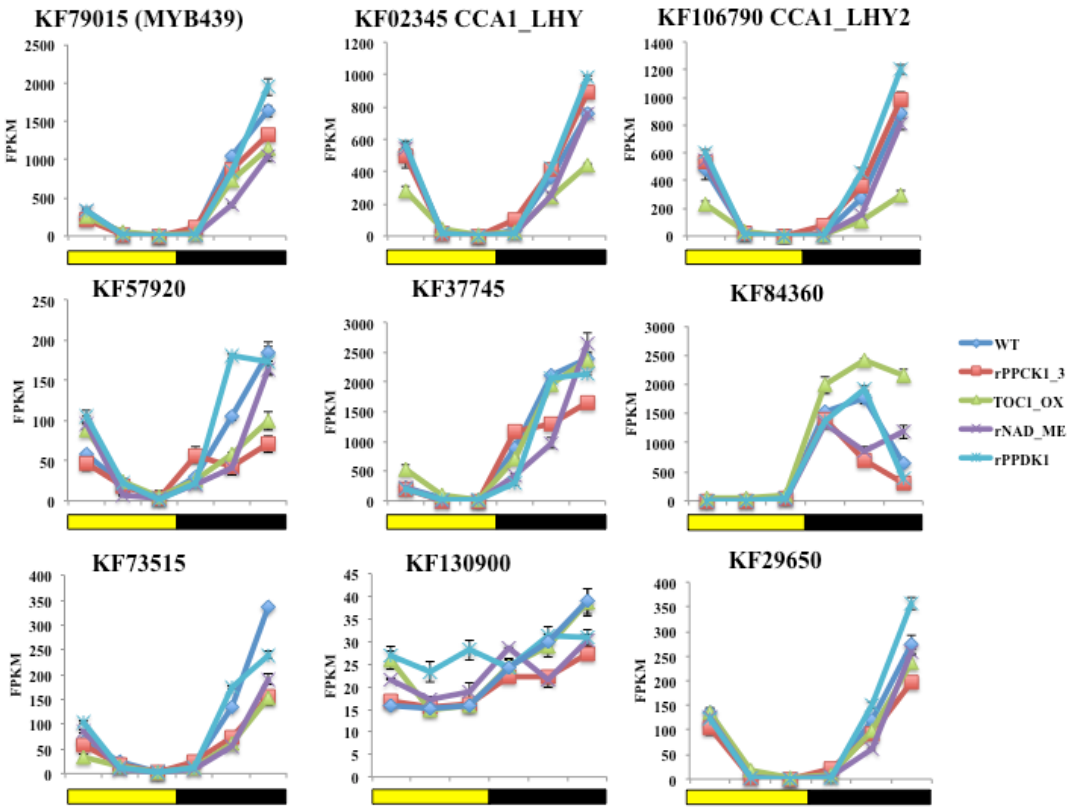

Figure 4.5. –Data collected from sequencing transgenic lines, both overexpressing: *TOC1_OX*, and RNAi knockdown lines: *rPPCK1_3*, *rNAD-ME* and *rPPDK*, showing changes in transcript regulation of the *Kalanchoe fedtschenkoi* RVE family. *TOC1* overexpression can be seen to repress *CCA1/LHY* expression, which is previously known.

As mentioned in Chapter 3, a large number of genes identified due to their induction and light/ dark regulation in the global RNA-seq analysis of C_3 and CAM leaves of *K. fedtschenkoi* were targeted in previous work for genetic manipulation in *K. fedtschenkoi.* Many of targeted genes possessed previously known functions in either the circadian clock, or CAM biochemistry. For some of these existing transgenic *K. fedtschenkoi* lines, quantitative RNA-seq analysis was performed by other members of the Hartwell lab using the Illumina Hi-Seq system (Boxall, Dever and Hartwell, unpublished data). This dataset allowed transcriptome-wide analysis of the regulation

of the majority of the genes in the genome when key CAM and clock genes were mis-regulated due to being targeted by a transgene. CAM leaves (LP6) were sampled every 4 h over a 12:12 light/ dark cycle from wild type plus a *KfPPCK1* RNAi knockdown line (*rPPCK1_3*), an iceplant *McTOC1* full-length over-expressor line (*TOC1_OX*), a *KfNAD_ME_b1* RNAi knockdown line (*rNAD_ME*), and a *KfPPDK* RNAi knockdown line (*rPPDK1*).

As with the previous SOLiD RNA-seq quantitative data for wild type leaves, KF130900 was again found to be the lowest abundance transcript in the RVE family within CAM leaves only reaching 40 FPKM at its peak (Fig. 4.5). KF130900 also displayed a low amplitude oscillation over the light/ dark cycle compared to the other RVE family members, with only a relatively small increase in transcript abundance at its peak at 22:00, 2 h before dawn. By contrast, the next lowest abundance *RVE* gene, KF57920, reached an FPKM value of 173 at its peak at 22:00, and the most abundant *RVE* gene (KF37745) reached a peak transcript level of 2381 FPKM (Fig. 4.5).

As the *RVE* family have been demonstrated to play a role in the circadian clock, plus it has also been shown that some play antagonistic functions to others in *Arabidopsis*, the results from the *McTOC1*-overexpressing line are of particular interest, especially as *TOC1* has been demonstrated to repress *CCA1/LHY* expression. A similar regulatory interaction can be seen for *K. fedtschenkoi* in Figure 4.4, where the two *K. fedtschenkoi CCA1/LHY* genes, KF106790 and KF02345, showed greatly reduced transcript levels in the *TOC1_OX* line at the time of their peak transcript level in the wild type, 22:00 (KF106790 reached 33 % and KF02345 reached 58 % of their wild type peak levels at 22:00 in the *TOC1_OX* line).

KF84360 was found to have increased transcript levels in the *TOC1_OX* lines, whilst KF37745 transcript levels in *TOC1_OX* were very similar to the wild type throughout

the light/ dark cycle (Fig. 4.5). All other *RVE* genes showed a reduced level of transcript in the *TOC1_OX* line relative to the wild type (Fig. 4.5). These findings supports the proposal that these genes are likely to carry out different functions, and are also regulated differentially within the circadian clock and its output pathways. In the *rPPCK1_3* line, which fails to phosphorylate PPC in the dark due to a failure to increase PPCK levels, and performs only 30 % of wild type dark CO_2 fixation (Boxall *et al*., 2015, in preparation), many of the RVE genes, including *KfMYB439*, displayed a decrease in transcript levels, especially at the time of the daily peak of that *RVE* gene in the wild type. However, the *CCA1/LHY* orthologs *KF106790* and *KF02345* displayed a slight increase in transcript levels relative to the wild type in line *rPPCK1_3*. In line *rNAD-ME*, which has less than 10 % of the wild type level of NAD-ME activity and performs little if any dark CO_2 fixation (Dever *et al*., 2015), again, several *RVE* genes had small reductions in their transcript levels relative to the wild type. The *CCA1/LHY* orthologs showed little difference from the wild type over the entire light/dark cycle. Finally, the *rPPDK* line, which possesses less than 10 % of the wild type level of PPDK activity and fails to fix CO_2 in the dark (Dever *et al*., 2015), had in general less of an impact on the transcript levels of the *RVE* genes compared to the other mutants. The *CCA1/LHY* orthologs, *KfMYB439, KF29650* and *KF57920* showed very small increases in transcript abundance in *rPPDK,* although theses differences were not greater than the standard errors suggesting they were not significant. Overall, the regulation of the *RVE* family in line *rPPDK* was very similar to the wild type over the entire light/dark cycle (Fig. 4.5).

4.2.2. The impact of the transgenic perturbation of *KfMYB439* levels on the light/dark regulation of key CAM- and circadian clock-associated genes.

As described in Chapter 3 the best RNAi and over-expresser lines for *KfMYB439* were identified. Detailed molecular, biochemical and physiological characterisation of phenotypes of these KfMYB439 transgenic lines was then conducted and is presented in this chapter.

As a first step towards understanding the phenotypic consequences of mis-expression of *KfMYB439* in transgenic lines of *K. fedtschenkoi*, semi-quantitative RT-PCR was used to investigate the regulation of a core set of known CAM-associated and circadian clock-associated genes in both the wild type and KfMYB439 lines using samples collected at 4 h intervals over a 12:12 light/dark cycle. The genes screened included the core clock genes *KfCCA1*(a single MYB-repeat TF in the RVE family) and *KfTOC1*, the CAM carboxylation pathway genes *KfGWD, KfMEX1, KfGPT2, KfPPCK1* and *KfPEPC*, and the CAM decarboxylation pathway genes *KfPPDK, KfPPDK-RP, KfNAD-ME α,* and *KfNAD- ME β.*

4.2.2.1 Light/ dark regulation of CAM and clock genes in *KfMYB439* full-length overexpresser lines

For the over-expression constructs, the CaMV35s promoter was used. Based on previous work with the *McTOC1* over-expression lines of *K. fedtschenkoi* (Dall'omo, 2011), these 35S promoter driven constructs were predicted to mediate constant over-expression of the full-length open reading frame of *KfMYB439* throughout the light/ dark cycle. In Chapter 3, initial phenotypic screens for the transcript level of *KfMYB439* in each line revealed that the transcript abundance of the gene was still under circadian control, despite the over-expression construct being under the control

of the 35S promoter (Fig. 3.2). The more detailed, complete LD timecourse presented here, which sampled CAM leaves (LP6) of wild and each *KfMYB439* transgenic line every 4 h over a 12:12 LD cycle starting at 2 h after lights on (02:00) confirmed this temporal regulation of the transcript abundance of *KfMYB439* in the over-expression lines (Figs. 4.6 & 4.7). The *KfMYB439* transcript abundance showed a light/ dark pattern of regulation similar to that in the WT, except that the over-expressing lines began to accumulate *KfMYB439* transcripts at the end of the light and had higher levels of *KfMYB439* than the WT throughout the dark period (Figs. 4.6 & 4.7). Whilst the daily trough of *KfMYB439* occurred at 14:00 (2 h into the dark period) in the WT, the trough shifted 8 h earlier to 06:00 (middle of the light period) in both KfMYB439 over-expresser lines. With these more detailed time course experiments, both *KfMYB439_FL* lines had similar levels of fold-increase in the transcript abundance of *KfMYB439*, in contrast to the results of the initial screening experiments reported in Chapter 3. Line *KfMYB439_FL_37C* had an average increase of 1.3-fold during the light and 1.6-fold during the dark (Fig. 4.6), whereas line *KfMYB439_FL_54B* showed a 1.5-fold increase during the light, and a 1.6-fold increase during the dark (Fig. 4.7). This suggests that there was a similar dosage effect in both lines, which is consistent with the uniform impact of the transgene on the other CAM and clock genes screened here (Figs. 4.6 & 4.7). From all of the genes whose light/dark pattern of regulation was investigated here, most showed little if any change in their steady-state transcript abundance over the whole of the light/ dark cycle (Figs. 4.6 and 4.7). Relative to the WT, the biggest differences in transcript abundance and temporal regulation were detected for *KfPPCK1,* and several genes involved in the central sugar and starch metabolism pathways (Figs. 4.6 and 4.67). *KfPPCK1* began to rise in abundance earlier in the dark period in both FL over-expression lines, and attained a

dark period peak 4 h ahead of the WT (Figs. 4.6 and 4.7). This earlier nocturnal rise and elevated peak of *KfPPCK1* transcript abundance was most marked in line *FL_37C* (Fig. 4.6). Of the genes associated with starch/ sugar metabolism and transport, *KfGPT2* (a *GLUCOSE 6-PHOSPHATE: PHOSPHATE TRANSLOCATOR* localised to the chloroplast envelope and hypothesised to be involved in daily starch synthesis and/ or nocturnal turnover) displayed a decrease in transcript abundance in line *KfMYB439_FL_54B* at 02:00, the time of its daily transcript peak in the WT (Fig. 4.7). In line *FL_37C*, a small decrease in *KfGWD* (*GLUCAN WATER DIKINASE),* which phosphorylates starch and plays a key role in starch breakdown in the dark, was detected at 22:00 (2 h before end of dark) (Fig. 4.6). Both *KfNAD_ME_β* (*β-subunit of NAD-ME*), and to a lesser extent *KfPEPC,* also showed small decrease in transcript abundance at 22:00 at the end of the dark in line *FL_37C* (Fig. 4.6). Overall, the only CAM- or clock-associated gene that showed a consistent change in transcript abundance in both FL over-expresser lines for *KfMYB439* was *KfPPCK1,* which was elevated relative to the WT at both 14:00 and 18:00 (early and mid-dark) in both of the *KfMYB439_FL* over-expresser lines studied here (Figs. 4.6 & 4.7).

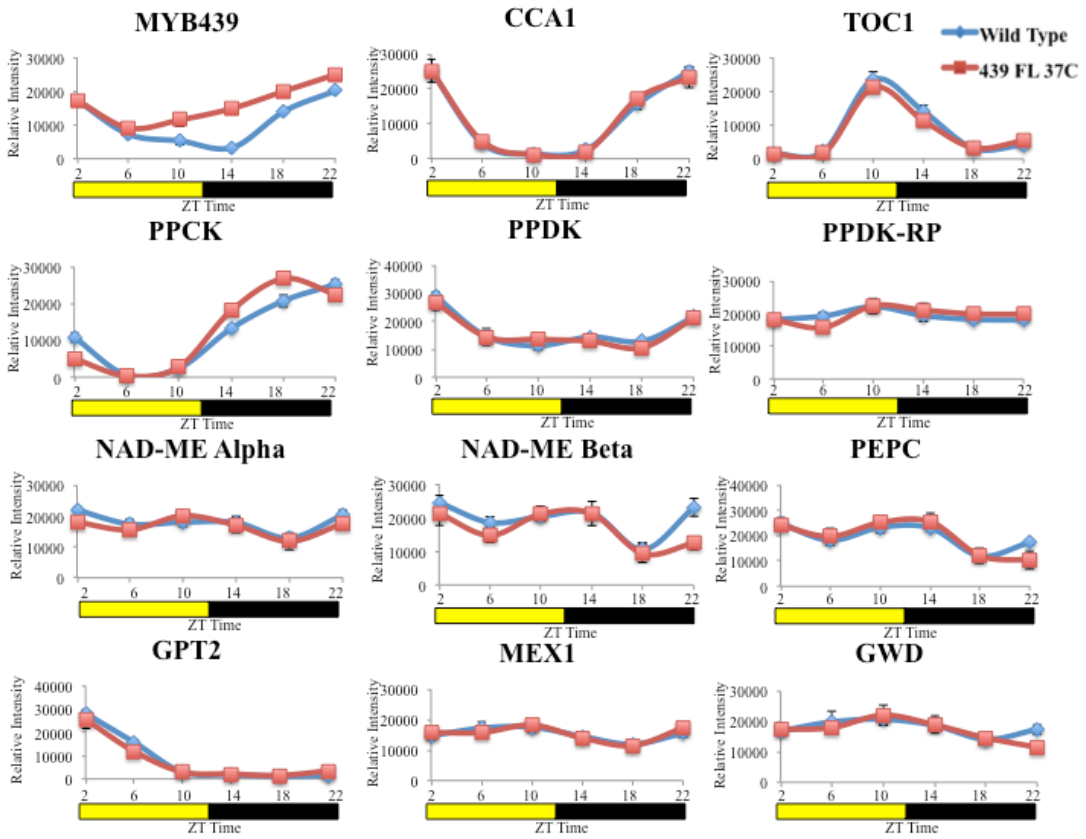

Figure 4.6. Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfMYB439_FL_37C* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line *FL_37C*. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

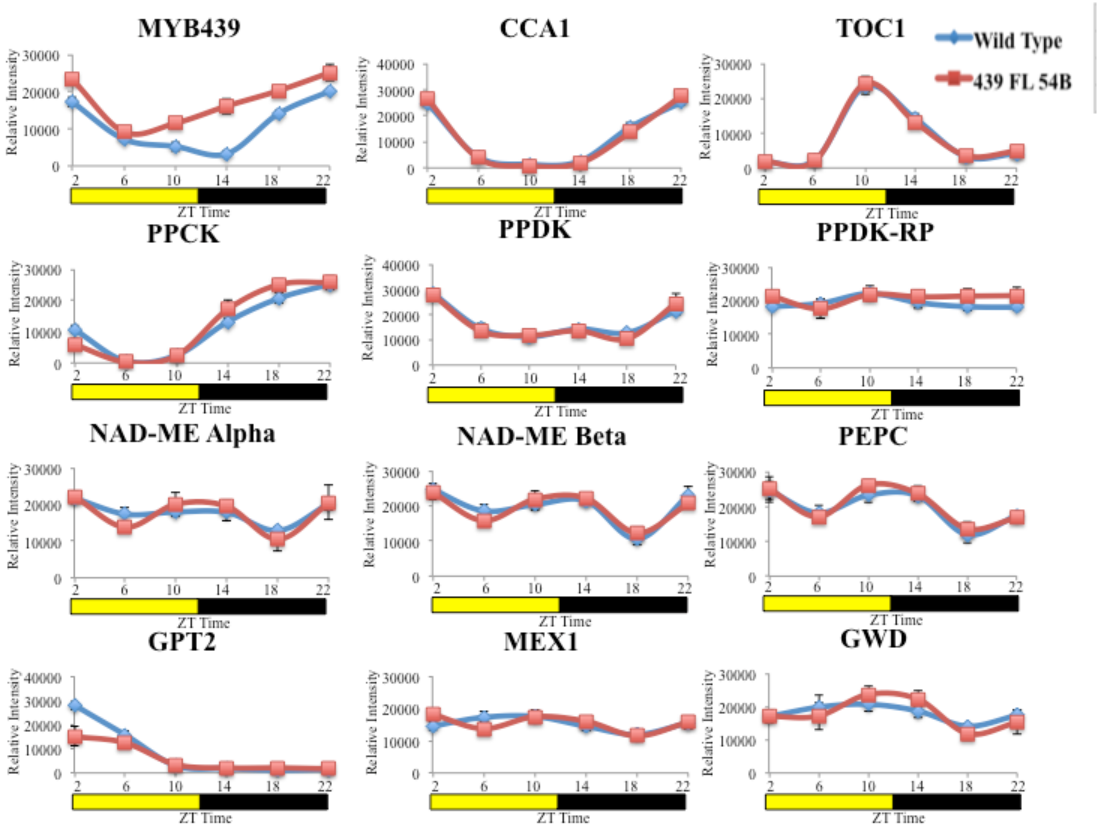

Figure 4.7. Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfMYB439_FL_54B* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line FL_54B. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

4.2.2.2 Light/ dark regulation of CAM and clock genes in *KfMYB439* RNAi knockdown lines

As described in Chapter 3, the fold-decrease in *KfMYB439* transcript levels in the RNAi lines was surprisingly low. This was unexpected, especially considering that RNAi has been estimated to lead to degradation of up to 80% of the endogenous transcript in Arabidopsis (Filichkin *et al.,* 2007). This supports the suggestion, already put forward in Chapter 3, that the high losses of the *KfMYB439* RNAi lines at both the tissue culture stage, and the soil establishment stage of the transformation process, may have been due to a lethal impact of strong silencing of *KfMYB439* on plant viability. Line *KfMYB439_RNAi_3C* showed a very slight decrease in the transcript level of the target transcript (1.1-fold decrease in the light and 1.02-fold decrease in the dark; Fig. 4.8). This does not support the data already presented for this line in Chapter 3 (Fig. 3.2), as there line 3C showed a 1.5-fold reduction at 22:00 at the end of the dark, whereas in the more detailed timecourse in figure 4.7, the decrease at 22:00 was only 1.1-fold (Fig. 4.8). *KfMYB439_RNAi_10B* showed a larger decrease of 1.4-fold in the light and 1.2-fold in the dark (Fig. 4.9), which was similar to the initial screening results, which indicated a 1.3-fold reduction (Fig. 3.2). These two lines were therefore expected to show dosage-dependent effects, as even the small differences in the extent of *KfMYB439* silencing may allow insights into the functions and downstream targets of *KfMYB439*.

The genes that showed the largest knock-on perturbations in transcript levels in these *KfMYB439_RNAi* lines again included *KfPPCK1,* and several genes associated with starch and sugar metabolism and transport, including *KfGPT2* and *KfGWD*. *KfPPCK1* transcript levels again rose earlier in the dark period and peaked earlier, in the middle of the dark, whereas the WT peaked at 22:00, 2 h before dawn (Figs. 4.8 & 4.9). Both

the increase in *KfMYB439* in *FL_37C* and *FL_58B,* and the decrease in *KfMYB439* levels in *RNAi_3C* and *RNAi_10B,* led to the same pattern of change in *KfPPCK1* levels.

KfGPT2 levels increased slightly at the post-dawn 02:00 transcript peak in line *RNAi_3C* (Fig. 4.8), but decreased in line *RNAi_10B* (Fig.4.9). The *KfGPT2* decrease in line *RNAi_10B* at 02:00 was one of the largest fold-changes amongst the transcripts assayed, and was consistent with the stronger effect of the RNAi construct on *KfMYB439* levels in that line. It is noteworthy that *KfGPT2* transcript levels were reduced in both the *KfMYB439_FL_54B* over-expressor line and the knockdown line. Thus, despite the opposite impacts of the different transgene constructs on the transcript level of *KfMYB439*, the downstream impacts on *KfGPT2* levels were the same (Figs. 4.7 and 4.9).

KfGWD showed a small increase in transcript abundance in both RNAi lines during its dusk phased peak at 10:00 to 14:00 (Figs. 4.8 & 4.9). This increase was greatest in line *RNAi_10B* at 14:00, consistent with the greater impact of the RNAi construct on the level of the target *KfMYB439* transcript in this line (Fig. 4.9).

Also, there were small changes in *KfNAD-ME α* and *β,* and *KfPEPC*, especially at 02:00 (2 h into the light), in line *RNAi_10B,* where transcript levels of all three genes were reduced in the RNAi line relative to the WT (Fig. 4.9).

Finally, the central clock genes *KfCCA1* and *KfTOC1* showed only very minor changes in transcript abundance suggesting that the operation of the core loop of the circadian oscillator was not influenced by the small decreases in *KfMYB439* achieved in these RNAi lines (Figs. 4.8 & 4.9).

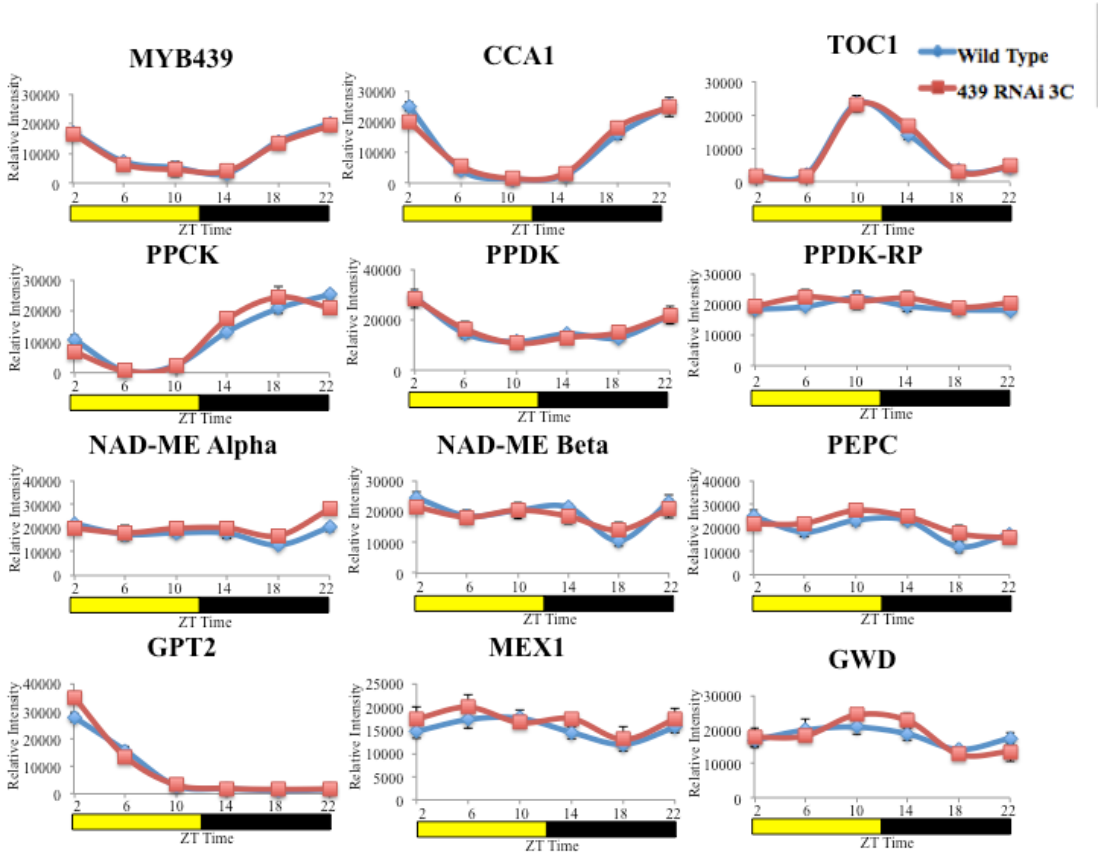

Figure 4.8. Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfMYB439_RNAi_3C* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line RNAi_3C. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

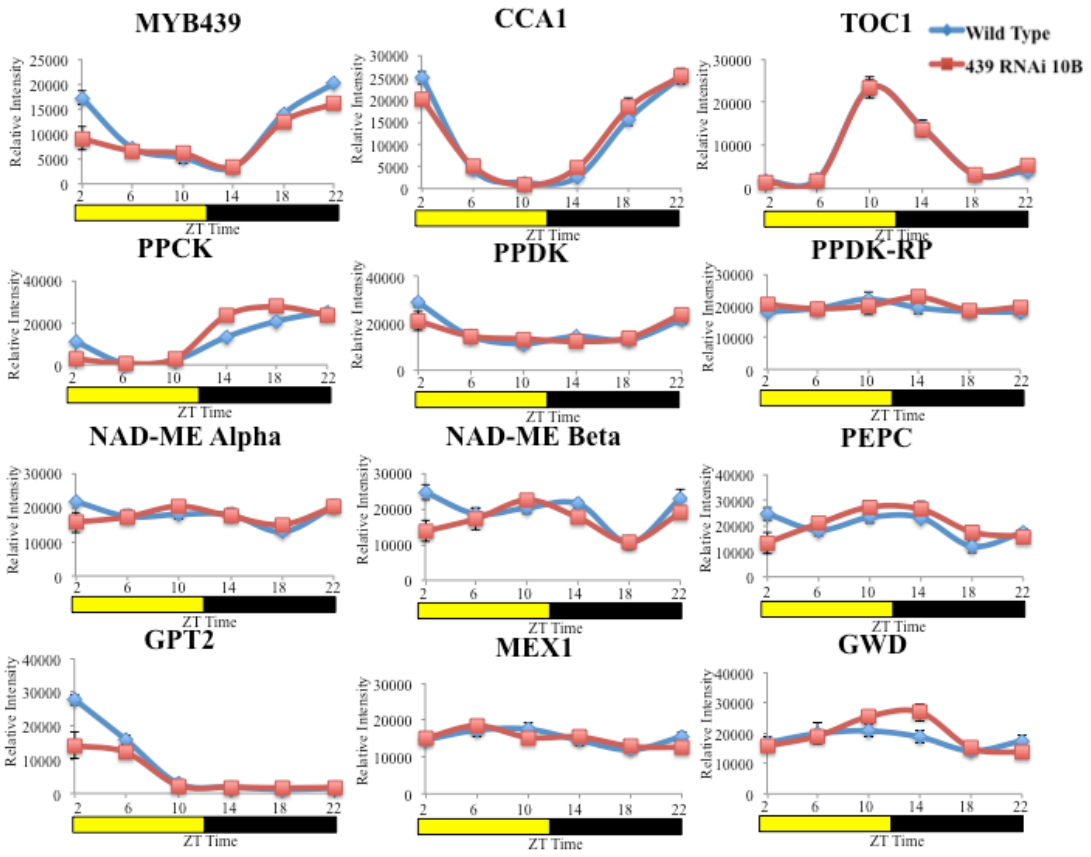

Figure 4.9. Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfMYB439_RNAi_10B* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line RNAi_10B. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

4.2.3. The impact of perturbing *KfMYB439* levels on circadian oscillations in the transcript abundance of circadian clock controlled genes

In Chapter 3 it was demonstrated that the transcript abundance of *KfMYB439* oscillated robustly under constant conditions (LL, constant light, constant temperature, constant humidity; Fig. 3.1). This demonstrated that this RVE family single MYB-repeat TF was under circadian clock control. Based on this finding, and the known role of several *RVE* genes in Arabidopsis in the operation of the core molecular oscillator, it was important to investigate whether the perturbations in *KfMYB439* levels achieved in the FL over-expresser and RNAi knockdown lines had a downstream impact on the circadian control of central clock genes and clock-controlled genes. One overexpresser, line *FL_37C* and one RNAi knockdown, line *RNAi_10B* were selected for a detailed LL time course experiment. Plants were grown from clonal plantlets in the greenhouse under 16:8 long days and then entrained for 1 week in 12:12 LD cycles in the Snijders Microclima growth cabinets. Samples were initially collected at 10:00 L and 22:00 D during the final light/dark cycle prior to release into LL constant conditions. The LL constant conditions commenced at dawn after the 22:00 D sample and the plants were not sampled for the first 24 h to ensure that all LL sampling occurred under truly free-running conditions. Samples of leaf pair 6 (CAM leaves) from WT and the transgenic lines were collected at 4 h intervals starting a 26:00 h LL (2 h into the subjective light period).

4.2.3.1 The impact of perturbing *KfMYB439* levels on circadian oscillations in *KfMYB439* full-length overexpression lines

In LL constant conditions, the *KfMYB439* FL over-expresser line showed a very similar pattern of *KfMYB439* regulation to the WT, but it had a consistently higher level of *KfMYB439* transcripts than the WT throughout the LL period, in accordance with its over-expression of *KfMYB439* from the 35S promoter driven transgene construct (Fig. 4.10). When the *KfMYB439* transcript rose to a high level during the third subjective day and night (Fig. 4.10; second white and black dotted boxes) it correlated with a knock-on effect on the central clock genes *KfCCA1* and *KfTOC1* which both failed to reach levels as high as those in the WT during that section of the LL experiment. This therefore suggests that very high levels of *KfMYB439* were able to affect the core clock. *KfPPCK* also followed a similar pattern to the WT, but did have lower peak levels of transcript in LL for three of the four rhythmic LL peaks (Fig. 4.10), whereas in LD the peak transcript level for *KfPPCK1* was higher and peaked earlier (Fig. 4.6).

Importantly, these results demonstrate robust circadian oscillations of the CAM-associated, clock output pathway gene *KfPPCK* in line *FL_37C* (Fig. 4.10). This reveals that the level of over-expression of *KfMYB439* achieved in this line was not sufficient to cause arrhythmia of the core molecular oscillator, or of the key clock-controlled CAM regulatory gene *KfPPCK* under constant light free-running conditions. Aside from the relatively small changes in amplitude of the rhythms of the measured clock-controlled genes under LL conditions, the phase and period of the oscillations were very similar to the WT for line *FL_37C*.

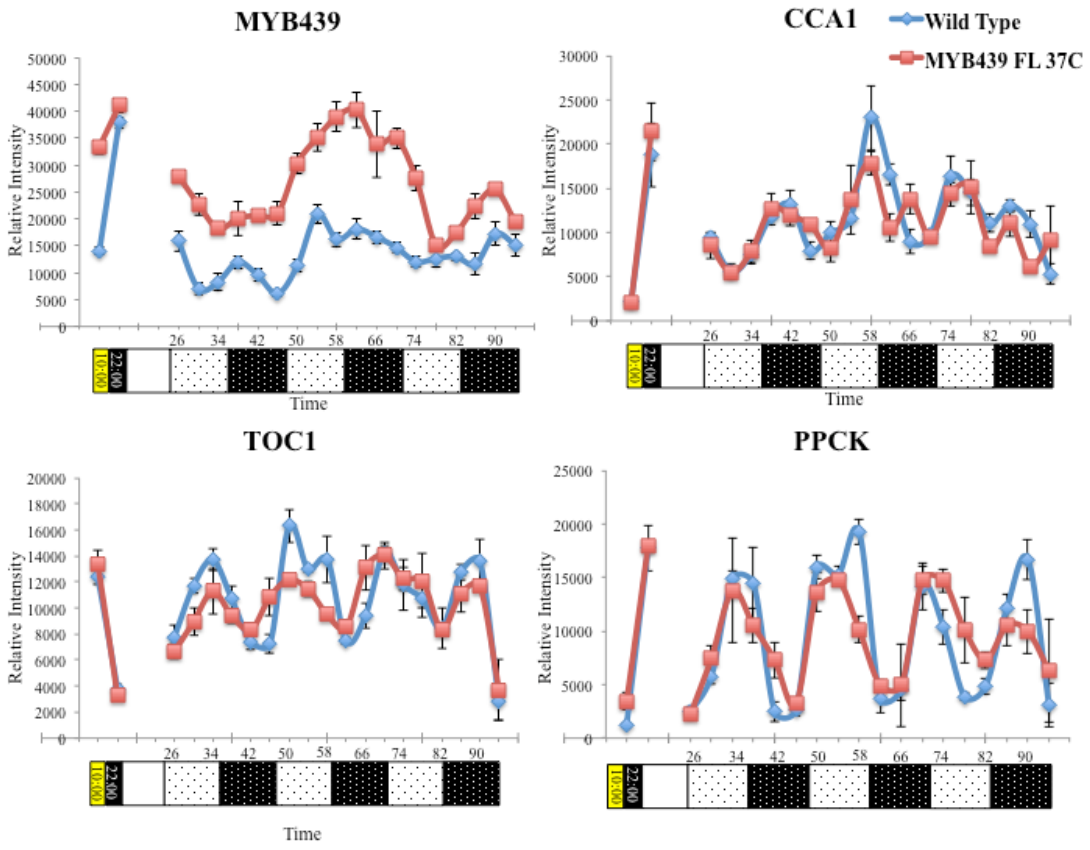

Figure 4.10. Constant light time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (Blue) and transgenic line *KfMYB439_FL_37C* (Red) measured in full CAM leaves (LP6) sampled at 4h intervals over a LL light regime. The first two samples collected were in LD and then the light regime was turned to LL. Plants were left for 26h in LL before sampling commenced. Fluctuations in the transcript abundance of clock genes (*KfCCA1* and *KfTOC1*) and CAM gene *KfPPCK* were determined, along with *KfMYB439* expression, for both Wild type and *KfMYB439_FL_37C*. The yellow bar represents light period, the black bar represents the dark period, and the white bar represents a break in sampling for the first 26h in LL. White boxes containing black dots represent subjective light and black boxes with white dots represent subjective dark. To entrain *KfMYB439_FL_37C*, The plants were entrained to lights on at ZT0 (00:00 h), and went off at 12:00h. Light sample was collected at 10:00 and dark sample was collected at 22:00. LL samples were collected from 26:00 in LL every 4 h for three days.

4.2.3.2 The impact of perturbing *KfMYB439* levels on circadian oscillations in *KfMYB439* RNAi knockdown lines

The reduction of the transcript level for *KfMYB439* in LL in the RNAi line 10B was not maintained consistently during the LL experiment (Fig. 4.11). After 42h LL, end of the second subjective dark period (Fig. 4.11; first black dotted box), the RNAi line peaked when the wild type level was at its minimum. This suggests that, as well as RNA degradation not being carried out properly in response to the RNAi construct, the circadian regulation of the transcript abundance has also been strongly perturbed in this line. At this same time point, 42 h LL, both *KfCCA1* and *KfTOC1* transcript levels were also very different to WT levels. *KfCCA1* transcript levels were much higher than in the WT, and *KfTOC1* levels were much lower. This analysis did not reveal large perturbations in the LL oscillations for *KfPPCK* compared to WT (Fig. 4.11).

Again, a key findings from these results is that the circadian clock continued to oscillate robustly in the *KfMYB439_RNAi_10B* line, for clock-output gene *KfPPCK* (Fig. 4.11). In both the *FL_37C* line and the *RNAi_10B* line, the oscillations of *KfMYB439* itself did not display robust rhythms comparable to those observed originally for CAM leaves of *K. fedtschenkoi* (Fig. 3.1). It is unclear why this gene did not display the same robust oscillations in this experiment as those presented in Fig. 3.1. The wild type plants were clonally propagated from the same stock and the growth conditions used were identical between the two experiments. One key possibility is that the plants were raised in the greenhouse during different seasons of the year, prior to their entrainment to 12:12 LD cycles in the Snijders growth cabinet. As the quantity and quality of natural daylight hours varies throughout the seasons in Liverpool, it is likely that plants raised at different times of year in the greenhouse

under the supplemented 16:8 LD conditions will have subtle differences in their core components of the circadian clock. However, the rhythms of *KfPPCK* were robust in both experiments, so it is clear that the lack of robust rhythmicity in this experiment for *KfMYB439* was a specific effect (Figs. 4.10 & 4.11).

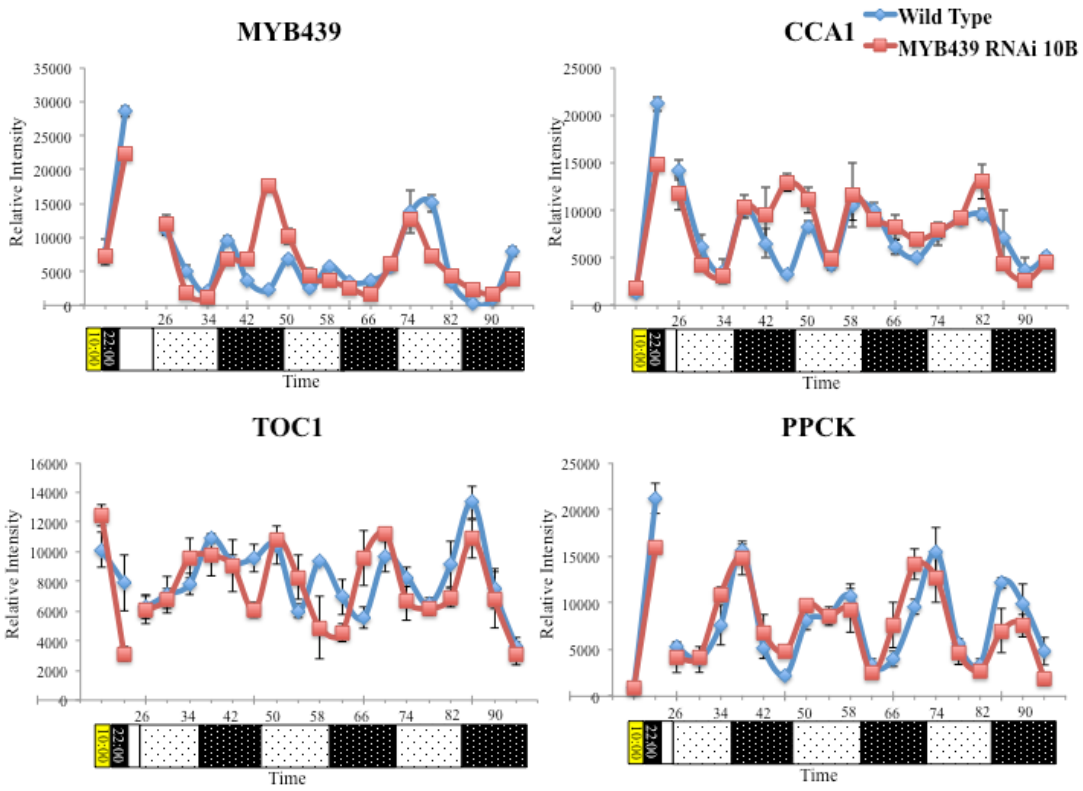

Figure 4.11. Constant light time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (Blue) and transgenic line *KfMYB439_RNAi_10B* (Red) measured in full CAM leaves (LP6) sampled at 4h intervals over a LL light regime. The first two samples collected were in LD and then the light regime was turned to LL. Plants were left for 26h in LL before sampling commenced. Fluctuations in the transcript abundance of clock genes (*KfCCA1* and *KfTOC1*) and CAM gene *KfPPCK* were determined, along with *KfMYB439* expression, for both Wild type and *KfMYB439_RNAi_10B*. The yellow bar represents light period, the black bar represents the dark period, and the white bar represents a break in sampling for the first 26h in LL. White boxes containing black dots represent subjective light and black boxes with white dots represent subjective dark. To entrain *KfMYB439_RNAi_10B*, The plants were entrained to lights on at ZT0 (00:00 h), and went off at 12:00h. Light sample was collected at 10:00 and dark sample was collected at 22:00. LL samples were collected from 26:00 in LL every 4 h for three days.

4.2.4 Effect of expression change of *KfMYB439* on *K. fedtschenkoi* gas exchange rhythms

All *KfMYB439* transgenic lines had their gas exchange rhythms measured using a multi-chamber Infra-Red Gas Analyser (IRGA) (Dever *et al*., 2015). Mature, developmentally synchronised plants which had been raised side-by-side in the greenhouse were entrained for 7 days in 12:12 LD conditions in the Snijders growth cabinet before sampling leaf pair 6 for the gas exchange measurements. Detached leaves had their petioles submerged in water immediately and parafilm was used to seal the water in the pots such that only the gas exchange by the leaf would be measured within the gas exchange chamber. Gas exchange measurements were recorded for at least 5 days in LD, before the growth cabinet settings were switched to LL conditions, and the leaves were measured for another 7 days. Only the latter 4 days of the LD segment of the experiment are illustrated, as the first 24 h after placing the leaves into the chamber are discarded as it takes this long for the gas exchange system to settle down after the start of the experimental run causing data in the first 24 h to be noisy and variable.

4.2.4.1 Impact of overexpressing *KfMYB439* on diurnal gas exchange rhythms.

Both *KfMYB439_FL* over-expresser lines displayed very different gas exchange patterns relative to the wild type, in both LD and LL conditions (Figs. 4.12 & 4.13). Both lines showed clear alterations in their circadian control of the CO_2 exchange rhythm in LL relative to the WT. The large peak for dark CO_2 fixation, which normally occurred early in the dark period in the WT, started prior to dusk in line *KfMYB439_FL_37C*, and began much later and lasted into the beginning of the light period in line *KfMYB439_FL_54B*. This had the consequence that the average amount

of light-period CO_2 fixation was greater in the two overexpresser lines (Fig. 4.12 A and B; Fig. 4.13 A and B; 37C achieved +1753 % and 54B achieved +794 % of wild type levels of CO_2 fixation in the light. These results are an average of all of the data collected in LD). Both lines fixed less CO_2 than the WT during the dark period (Table 4.4; 37C achieved -58%, and 54B achieved -60% of WT dark CO_2 fixation). Following the change to LL conditions, the rhythms of CO_2 exchange are immediately different between the WT and the transgenic over-expresser lines. The WT displayed the well-defined, short-period CO_2 exchange rhythm characteristic of the CAM circadian rhythm in *K. fedtschenkoi* under LL conditions (Anderson and Wilkins, 1989; Dever et al., 2015). Within 36 h of the start of the LL conditions, when the wild type was fixing no CO_2, *KfMYB439_FL* CO_2 fixation was close to peaking in both lines (Fig. 4.12 C and Fig. 4.13 C). This emphasises that the period of the rhythm in both over-expresser lines is shorter than the WT period. At the end of the seven days in LL, the rhythms were truly anti-phasic, with the over-expresser lines achieving peak CO_2 fixation at times when the WT reached its trough of CO_2 fixation (Figs. 4.12 and 4.13). *KfMYB439_FL_37C* maintained higher amplitude and more robust rhythms than line 54B, which may be linked to the slightly lower over-expression level of the *KfMYB439* transcript for line 37C.

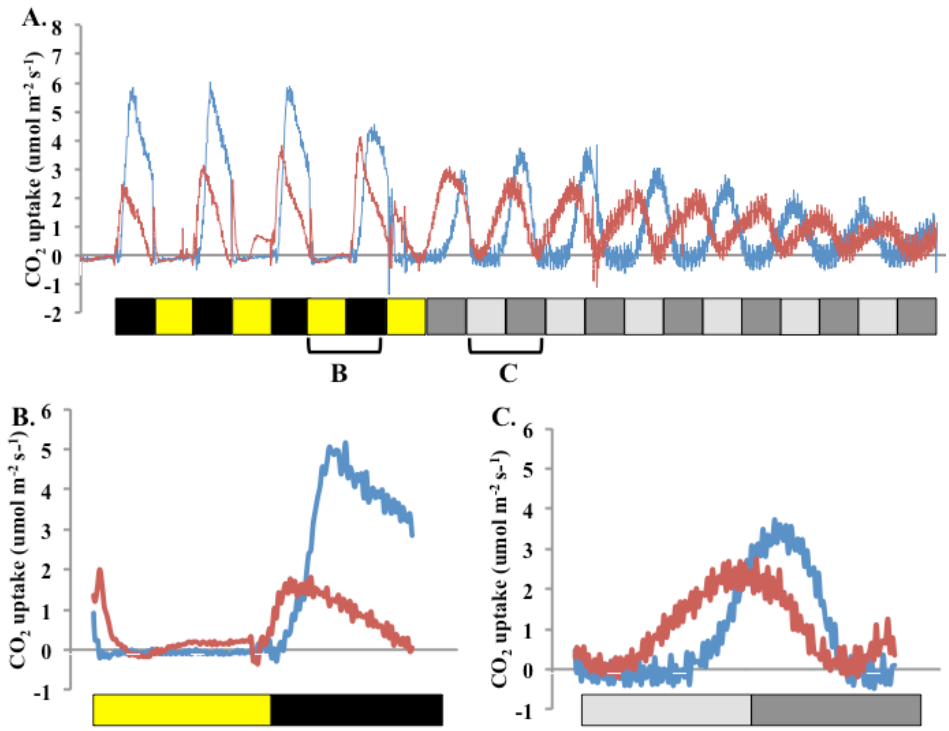

Figure 4.12. Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfMYB439_FL_37C*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfMYB439_FL_37C* is in red **A.** CO_2 fixation over 4 days LD followed by 7 days in LL. **B.** Snapshot of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C.** CO_2 fixation after entering LL, within 36h the plant is clearly already displaying a shorter period than WT. The selected 24 h window is indicated by the C symbol below graph A.

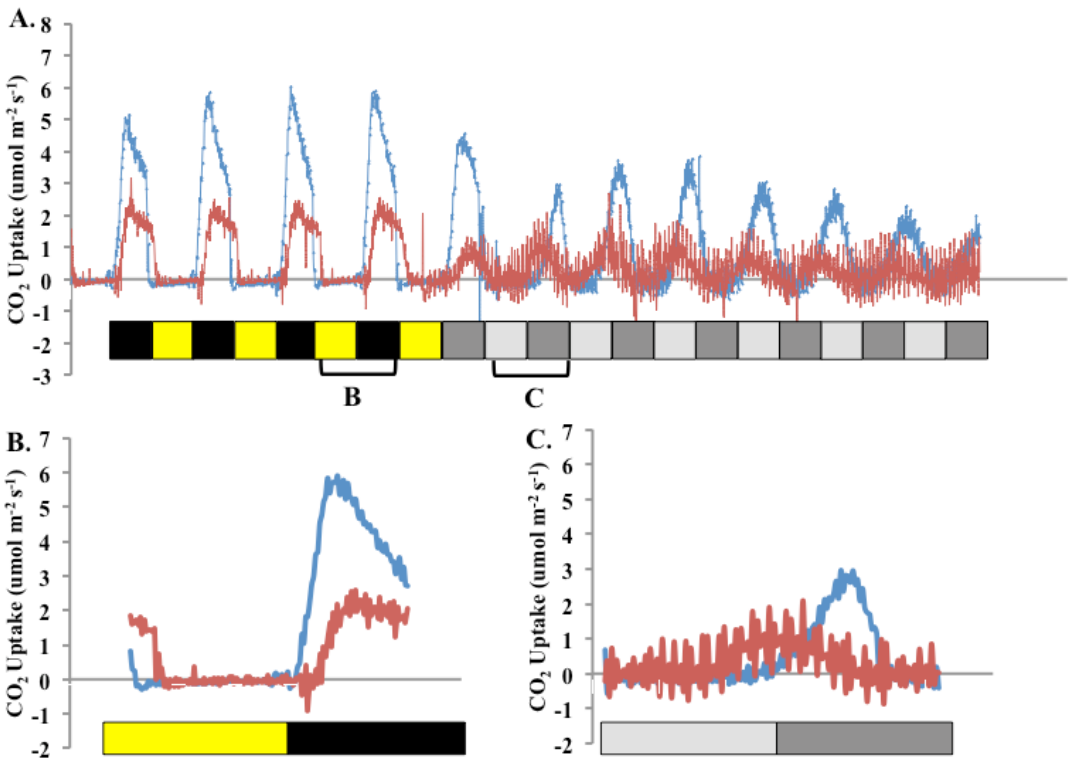

Figure 4.13. Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfMYB439_FL_54B*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfMYB439_FL_54B* is in red **A.** CO_2 fixation over 4 days LD followed by 7 days in LL**. B.** Snapshot of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C.** CO_2 fixation after entering LL, within 36h the plant is clearly already displaying a shorter period than WT. The selected 24 h window is indicated by the C symbol below graph A.

Table 4.4. Average net 24 h CO_2 fixation for each transgenic line and the WT over the entire light or dark period. Both transgenic lines fixed more CO_2 than the WT in the light, but less in the dark.

Line	Total CO_2 fixation over the 12 h light period (µmol CO_2 fixed $m^{-2)}$	% change	Total CO_2 fixation over the 12 h dark period (µmol CO_2 fixed $m^{-2)}$	% change
WT	-1760.4	-	140285.02	-
439 FL 37C	29101.52	1753	58885.95	-58
WT	-1360.65	-	122328.77	-
439 FL 54B	9436.95	794	48471.98	-60

4.2.4.2 Impact of RNAi knockdown *KfMYB439* lines on diurnal gas exchange rhythms

During the 12:12 LD cycles, *KfMYB439_RNAi_3C*, showed a very similar pattern of gas exchange to the wild type, although it did achieve substantially less dark CO_2 fixation overall than the WT (Fig. 4.14, Table 4.5). Interestingly, when the average net CO_2 exchange was calculated for the entire light period, line *RNAi_3C* lost even more respiratory CO_2 than the WT in the light (Fig. 4.14, Table 4.5). *KfMYB439_RNAi_10B* achieved a reduced level of dark CO_2 fixation relative to the WT under LD cycles, and also possessed a pronounced phase II (short burst of CO_2 fixation in the few hours after dawn) (Fig. 4.15). This phase II CO_2 fixation was the major contributor to line *RNAi_10B* achieving 453 % greater CO_2 fixation than the WT during the light period (Fig. 4.15, Table 4.5). Both RNAi lines fixed less CO_2 during the dark period, but this was again less dramatic in line *RNAi_3C* (3C = -19%, 10B = -68%; Table 5). Following the transition from LD cycles to LL free-running conditions, both RNAi lines displayed a similar circadian phenotype to the over-expresser lines (Figs. 4.14 and 4.15). Specifically, both lines had a phase shift in LL, where the RNAi lines consistently peak earlier than the WT under LL conditions, yet also appear to fix CO_2 for a longer period of time. The rhythm is also lower in amplitude that dampens more rapidly as the LL time course proceeded (Fig. 4.14, Fig. 4.15). In the stronger RNAi line 10B, the rhythm dampened to almost complete arrhythmia by the end of the 7-day experimental run in LL. The first peak in LL is earlier in 10B than the wild type, and also the duration of each period of net CO_2 fixation under LL conditions (period from one trough to the next) was longer in the two RNAi lines (Figs. 4.14C and 4.15C). These results suggest that *KfMYB439* is

likely to play a role in optimising the circadian clock, and/ or outputs from the circadian clock.

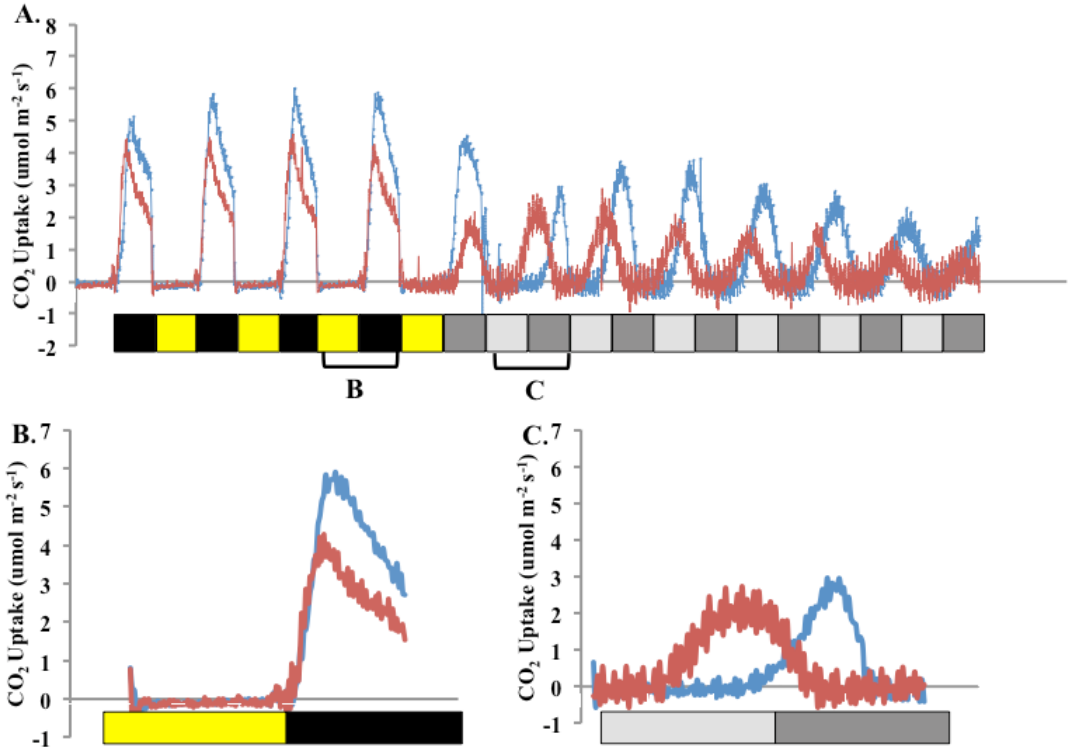

Figure 4.14. Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfMYB439_RNAi_3C*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfMYB439_RNAi_3C* is in red **A.** CO_2 fixation over 4 days LD followed by 7 days in LL. **B.** Snapshot of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C.** CO_2 fixation after entering LL, within 36h the plant is clearly already displaying a shorter period than WT. The selected 24 h window is indicated by the C symbol below graph A.

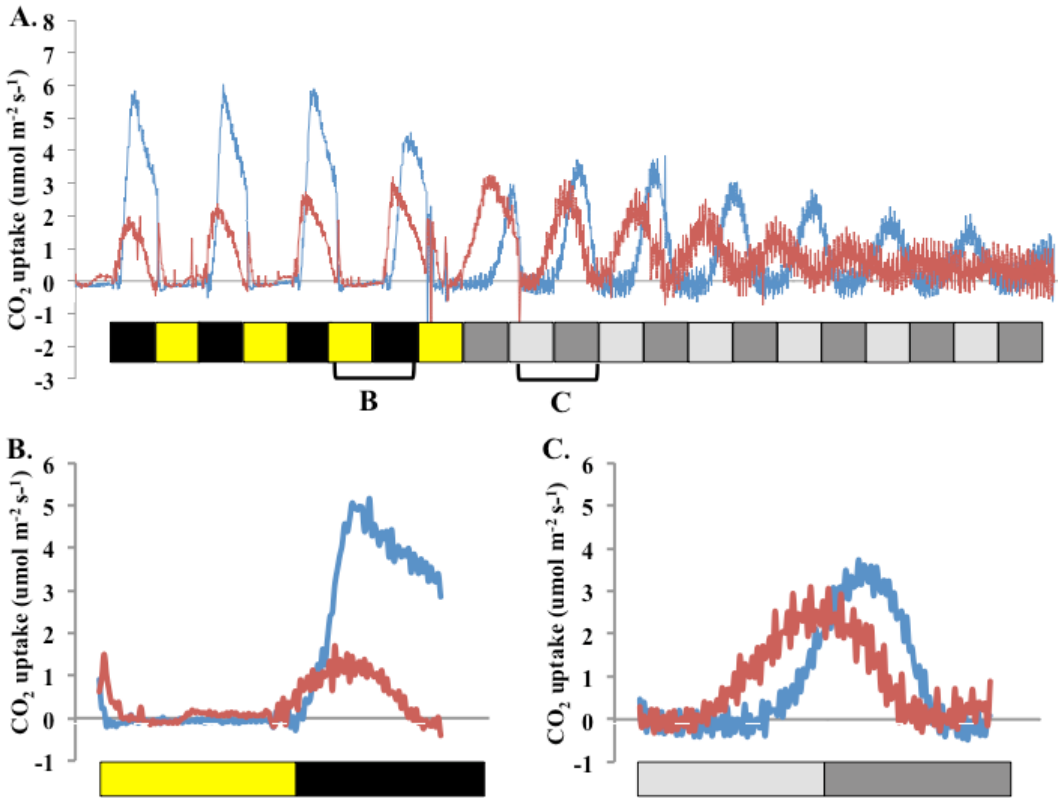

Figure 4.15. Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfMYB439_RNAi_10B*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfMYB439_RNAi_10B* is in red **A.** CO_2 fixation over 4 days LD followed by 7 days in LL. **B.** Snapshot of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C.** CO_2 fixation after entering LL, within 36h the plant is clearly already displaying a shorter period than WT. The selected 24 h window is indicated by the C symbol below graph A.

Table 4.5. Net amount of CO_2 fixed by each line and their associated wild types.

Line	Total CO_2 fixation over the 12 h light period (µmol CO_2 fixed $m^{-2)}$	% change	Total CO_2 fixation over the 12 h dark period (µmol CO_2 fixed $m^{-2)}$	% change
WT	-1360.65	0	122328.77	0
439 RNAi 3C	-4476.98	-229	99194.1	-19
WT	-1760.4	0	140285.02	0
439 RNAi 10B	6212.6	453	44295.83	-68

4.2.5 Measurements of the end products of CAM using metabolite assays

4.2.5.1 Impact of *KfMYB439* Full-Length Overexpression lines on metabolite levels

4.2.5.1.1 Daily fluctuations in malate

In Chapter 3, the pH indicator dye, chlorophenol red was used to estimate pH and therefore cell sap acidity, which, in CAM species, is assumed to be largely a direct result of malate accumulation and turnover. This assay allows for a crude estimation of the amount of malate present. However, it was important to quantify the amount of malate produced by the end of the dark period, and turned over by the end of the light period, more accurately. Quantitative measurements of malate were therefore made on dawn and dusk hot methanol soluble metabolite extracts of CAM leaves from both the WT and the transgenic lines using a well-established enzyme-linked spectrophotometric assay for malate (Mollering 1974).

Both the WT and the two *KfMYB439_FL* over-expresser lines displayed significant fluctuations in malate between dawn (high malate) and dusk (low malate) (Fig. 4.16) The mean differential between the dawn and dusk level of malate (Δ-malate = malate at dawn – malate at dusk) was 40.1 μmol gFW^{-1} for the WT but 26.6 μmol gFW^{-1} for *FL_37C* and 18.7 μmol gFW^{-1} for *FL_58B* (Fig. 4.16). Thus, the over-expressing lines displayed, on average, a lower level of malate accumulation during the dark period, and a higher retention of malate at the end of the light period, relative to the wild type (Fig. 4.16) However, due to the larger variations between the values obtained for the biological replicates for the over-expressor lines, the differences between nocturnal accumulation, light period turnover, and daily Δ-malate relative to the WT could not be confirmed as statistically significant (Fig. 4.16).

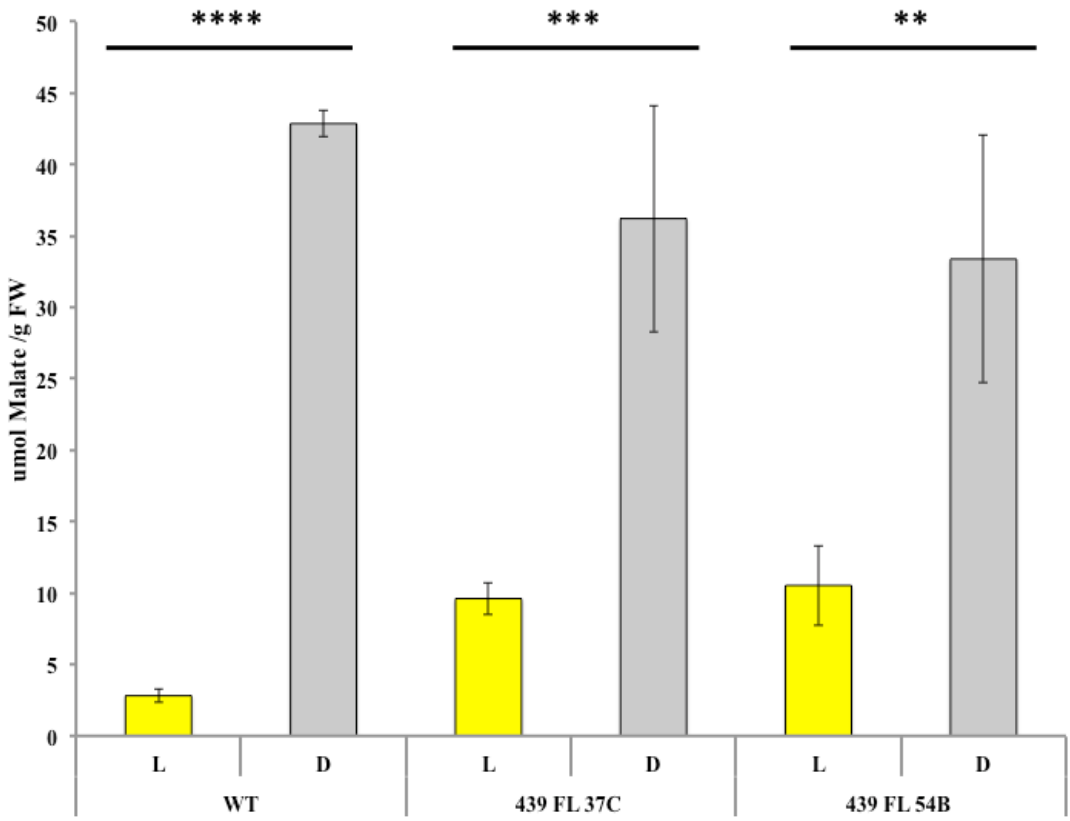

Figure 4.16. Diurnal fluctuations in malate concentrations in *KfMYB439_FL_37C* and *KfMYB439_FL_54B*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Malate levels were low at the end of the light period due to malate decarboxylation occurring during the day, and high at the end of the dark due to nocturnal CO_2 fixation resulting in malic acid accumulation. Malate levels for *KfMYB439_FL* lines were higher than WT at the end of the light period and malate levels were slightly lower at the end of the dark. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $\leq$ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences relative to the corresponding WT value.

4.2.5.1.2 Daily fluctuations in leaf starch

Malate and starch levels in CAM leaves reciprocate over the light/ dark cycle. In chapter 3, only a visual iodine staining technique was used to investigate leaf starch levels in the transgenic lines, which provides at best a crude quantitative estimate of the level of starch in a leaf disc.

Relative to the WT, both over-expression lines had lower starch levels at the end of the dark (WT = 1.4 mg Starch gFW^{-1}, 37C = 0.3 mg Starch gFW^{-1}, 54B = 0.4 mg Starch gFW^{-1}), although only *KfMYB439_FL_37C* was calculated to have a statistically significant difference (Tukey HSD, P = 0.0389). During the light period, the over-expression lines displated contrasting results. Line *FL_37C* accumulated very similar levels of leaf starch to the WT in the light period (WT = 2.4 mg Starch gFW^{-1}, 37C = 2.5 mg Starch gFW^{-1}; Fig. 4.17), but *FL_54B* had greatly reduced levels (1.3 mg Starch gFW^{-1}), despite having similar levels of malate. Using these values to calculate the daily turnover of starch, whilst the WT used 1 mg starch gFW^{-1} per day, line *FL_37C* used 2.2 mg starch gFW^{-1} per day, and line FL_54B used 0.9 mg starch gFW^{-1} per day. Considering the reduced levels of dark CO_2 fixation and daily Δ-malate values in the FL lines relative to the WT, it is clear that the greater daily starch turnover in line *FL_37C* did not provide increased PEP for dark CO_2 fixation and malate accumulation. It is possible that line *FL_37C* exported more soluble sugars, derived from starch degradation, from the CAM leaves to fuel respiration and growth elsewhere in the plant. Alternatively, it may be that this line underwent greater levels of futile cycling within the metabolic pathways associated with CAM and thus burnt more of its starch without achieving a greater level of dark CO_2 fixation and malate accumulation. However, line *FL_54B* turned over a similar amount of starch to the WT and yet this line had a slightly smaller Δ-malate than line

FL_37C, and fixed less CO_2 in the dark period than line *FL_37C* (Fig. 4.16 and Table 4.4). Thus, there was a clear disconnect between the relative amounts of daily of starch turnover and the amount of CO_2 fixation and malate accumulation and turnover in the two FL over-expresser lines.

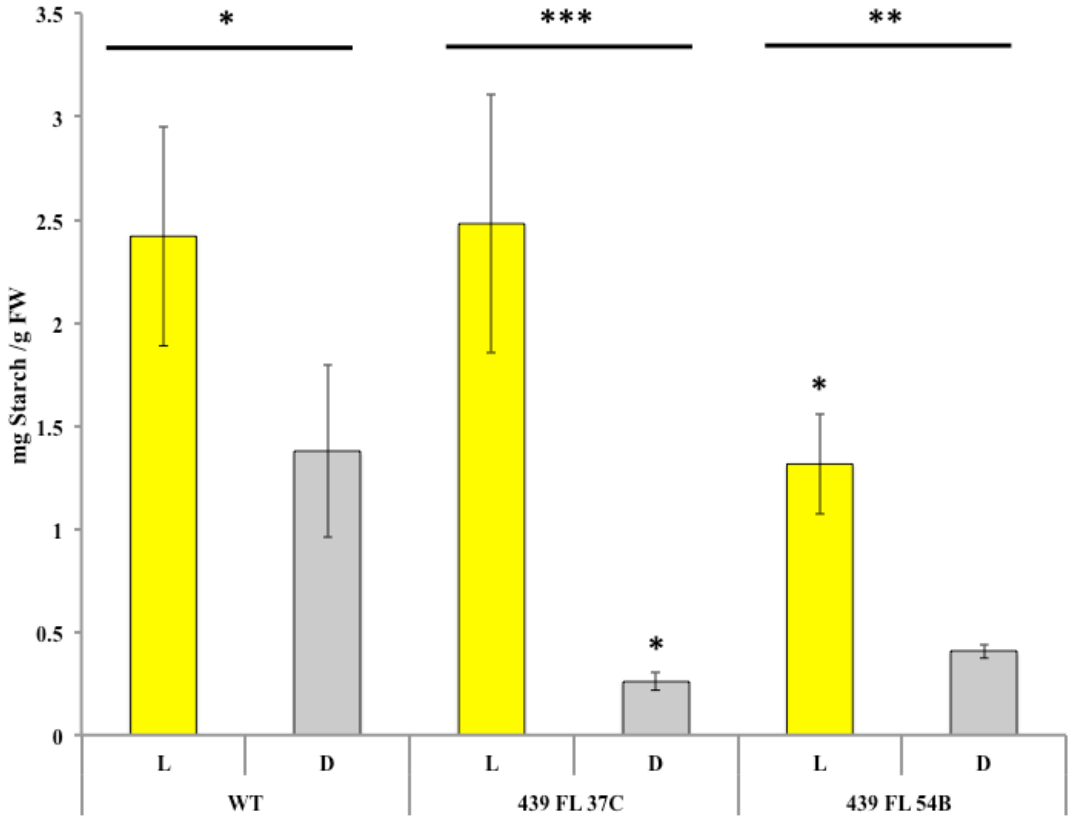

Figure 4.17. Diurnal fluctuations in starch concentrations in *KfMYB439 FL_37C* and *KfMYB439_FL_54B*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Starch levels are highest at the end of the light period, after photosynthesis has been occurring all day. Starch levels then drop to lower levels at the end of the dark, due to starch stores being mobilised to produce PEP for nocturnal CO_2 fixation. Both FL lines show an average reduction in starch levels during the dark period, whilst only *KfMYB439_FL_54B* also shows decreases in the light too. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

4.2.5.1.3 Daily fluctuations in soluble sugars

Wild type levels of glucose and fructose were higher at the end of the light than at the end of the dark, whilst the level of sucrose in the WT leaves was higher at the end of the dark than at the end of the light (Fig. 4.18). Both FL lines had a higher amount of glucose at the end of the light than the WT, but whilst line *FL_37C* had less glucose at the end of the dark than the WT, line *FL_54B* had a slightly higher end-of-dark mean glucose level than the WT, although this was not statistically significant (Fig. 4.18). Both FL lines had higher glucose levels at the end of the light than at the end of the dark, as seen in the WT (Fig. 4.18). The daily turnover of glucose between the end of the light and the end of the dark was higher than the WT for both FL lines, with line *FL_37C* displaying the greatest turnover of glucose between dusk and dawn.

There was a trend of elevated glucose levels at the end of the light period in the FL lines, with line *KfMYB439_FL_54B* accumulating significantly more, with over 3 times as much glucose as the WT (Fig. 4.18; Tukey HSD, $P < 0.0001$).

Opposite to the WT, line *FL_37C* had very low fructose at dusk and high fructose at dawn, whereas *FL_54B* fructose levels varied little between dusk and dawn (Fig. 4.18). Line *FL_54B* accumulated less fructose than the WT at both dusk and dawn, whereas line *FL_37C* had significantly less fructose than the WT at dusk ($P<0.0001$), and approximately double the WT level of fructose at dawn (Fig. 4.18).

Whilst WT sucrose levels were higher at dawn than at dusk, both transgenic lines had higher levels of sucrose at dusk than at dawn, reversing the dusk/ dawn variation in sucrose observed in the WT (Fig. 4.18). Leaves of both FL lines contained significantly more sucrose than the WT at the end of the light (37C: $P < 0.0001$, 54B: $P = 0.0005$), and both lines contained less sucrose than the WT at the end of the dark (Fig. 4.18).

By adding all of the values for glucose, fructose and sucrose it was possible to calculate the total level of soluble sugars in the leaves at the end of the light and dark and the total amount of sugar measured across the two time points (Fig. 4.19). The wild type plants had similar total sugar levels at the end of the light and dark, due to the previously discussed higher glucose and fructose during the day, and higher sucrose during the night (Fig. 4.19). Both *KfMYB439_FL* lines showed a significant increase in the total amounts of soluble sugars at the end of the light period (37C & 54B = $P < 0.0001$), and both lines had a lower level of sugars at the end of the dark period (37C: $P = 0.0071$, 54B: $P = 0.0061$). The summative total amount of soluble sugars from both the end of light and dark samples revealed similar mean levels of soluble sugars across both the WT and transgenic lines (Fig. 4.19). Whilst the trend suggested a slight decrease in the sum total sugar levels in *FL_37C,* and a further slight decrease in *FL_54B*, relative to the WT, the large error bars on the values for the FL lines prevent further interpretation of these differences.

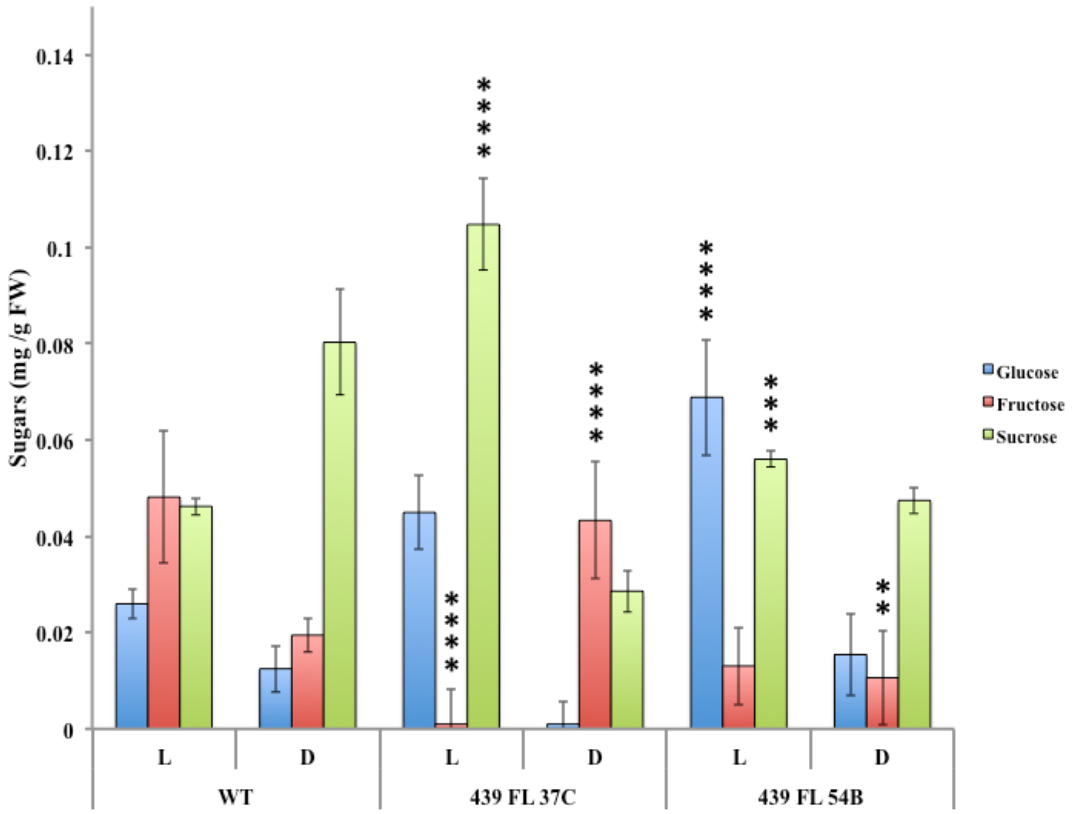

Figure 4.18. Levels of Glucose, Fructose and Sucrose in KfMYB439 overexpressing lines, at the end of the light (L; 23:00) and dark (D; 11:00) periods. *KfMYB439_FL_37C* shows differences in fructose levels during both L and D samples, and *KfMYB439_FL_54B* also shows low fructose at the end of the dark. *KfMYB439_FL_54B* also shows an increase in both glucose and sucrose in the light period. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $\leq$ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

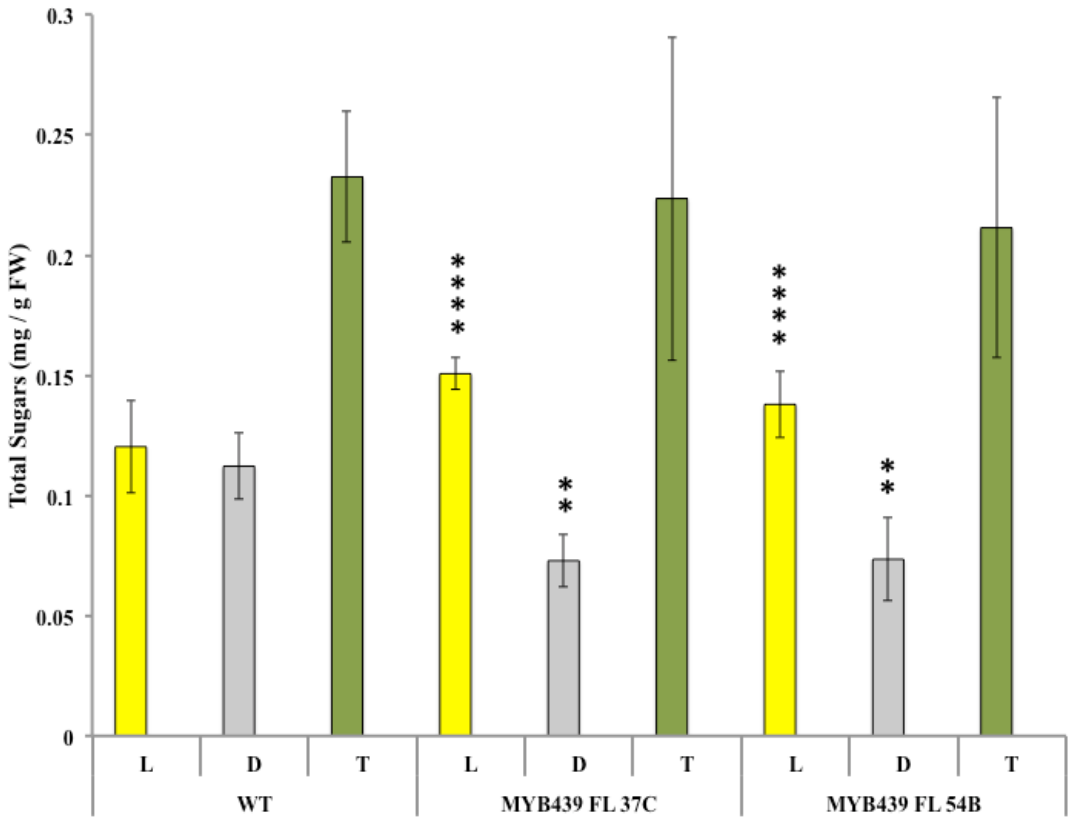

Figure 4.19. Variation in total soluble sugar levels in WT and the *KfMYB439 FL* lines. Yellow bars indicate light sugars levels (L), grey represent dark levels (D) and green bars show the total of sugar levels (T) calculated by summing the end-of-dark and end-of-light values for all sugars. There is a clear trend that these MYB439 FL lines have higher sugar levels during the day, and lower levels during the dark. Total sugar turnover in a day is approximately equal. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

4.2.5.2 Impact of *KfMYB439* RNAi Knockdown Lines on metabolite levels.

4.2.5.2.1 Daily fluctuations in malate levels

All three lines, WT and both *KfMYB439* RNAi lines, showed cycling of malate between dawn and dusk (Fig. 4.20). The *KfMYB439* RNAi lines both showed a reduction in malate accumulation compared to the WT and the end of the dark (3C: P =0.0066 , 10B: $p < 0.0001$). Both RNAi lines show significant cycling of malate between dawn (increased malate) and dusk (decreased malate). The Δ-malate values were 40.1 gFW^{-1} for wild type, 34.3 gFW^{-1} for *KfMYB439_RNAi_3C* and 18.7 gFW^{-1} for *KfMYB439_RNAi_10B*. Both lines show significant reductions in dark malate accumulation at the end of the dark period (3C: P =0.0066, 10B: $P < 0.0001$), and *KfYB439_RNAi_10B* also showed a significant increase at the end of the light period (P = 0.0038). *KfMYB439_RNAi_3C* also showed a slight increase in amount of malate in the light, but not enough to be significantly different.

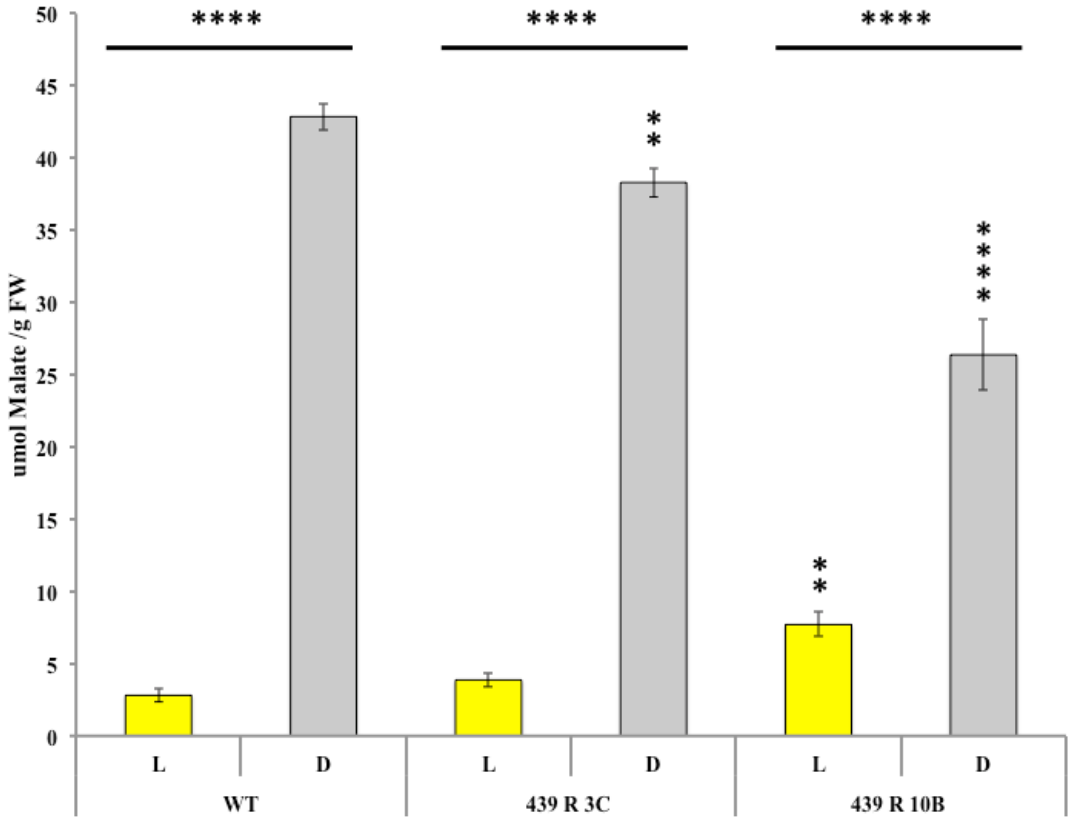

Figure 4.20. Diurnal fluctuations in malate concentrations in *KfMYB439_RNAi_3C* and *KfMYB439_RNAi_10B*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Malate levels were low at the end of the light period due to malate decarboxylation occurring during the day, and high at the end of the dark due to nocturnal CO_2 fixation resulting in malic acid accumulation. Malate levels for *KfMYB439_RNAi* lines were slightly higher than WT at the end of the light period and malate levels were lower at the end of the dark. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences relative to the corresponding WT value.

4.2.5.1.2 Daily fluctuations in starch

Starch also continues to cycle in these transgenic lines. Both RNAi lines have greatly reduced amount of starch by the end of the dark period compared to wild type (Fig. 4.21; 3C: P = 0.0059, 10B: P = 0.0045). It appears only small reductions in *MYB439* can affect starch breakdown during the dark. During the light period though, *KfMYB439_RNAi_3C* has similar levels to wild type, whereas *KfMYB439_RNAi_10B*, which has the bigger reduction in transcript, also shows a low level of starch during the day too (P = 0.0005). Therefore both malate and starch levels are significantly lower than wild type in *KfMYB439_RNAi_10B.* Wild type shows an average turnover of starch per day of 1mg. From calculating turnover of starch, *RNAi_3C* shows a great increase in turnover of 2.2 mg starch gFW^{-1} per day, whereas *RNAi_10B* shows a lower level of cycling, with only 0.7 mg starch gFW^{-1} per day.

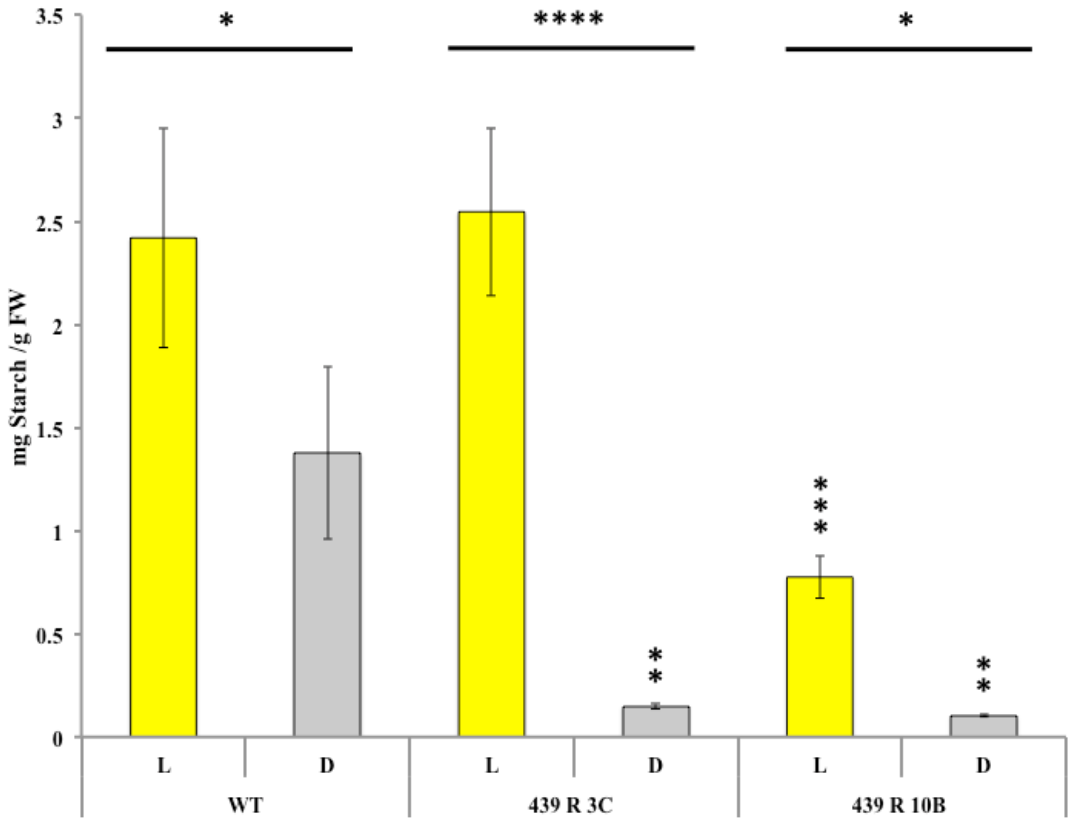

Figure 4.21. Diurnal fluctuations in starch concentrations in *KfMYB439_RNAi_3C* and *KfMYB439_RNAi_10B*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Starch levels are highest at the end of the light period, after photosynthesis has been occurring all day. Starch levels then drop to lower levels at the end of the dark, due to starch stores being mobilised to produce PEP for nocturnal CO_2 fixation. Both RNAi lines show a decrease in dark starch stores, and *RNAi_10B* also shows a decrease in end of light levels of starch, whilst *RNAi_3C* shows no difference to wild type. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

4.2.5.1.3 Daily fluctuations in soluble sugars

As with *KfMYB439_FL* lines, all these lines also show a glucose peak during the light period, and a drop at the end of the dark. For fructose, wild type shows a light time peak again, whereas *KfMYB439_RNAi_3C* appears to not cycle fructose, keeps a constant level, and *KfMYB439_RNAi_10B* has a large fructose peak at the end of the dark, which is much higher than seen in WT ($P < 0.0001$;Fig. 4.22). Again for sucrose the only major difference to wild type is for *KfMYB439_RNAi_10B* with sucrose at the end of the dark, which shows a much bigger peak than wild type ($P < 0.0001$).

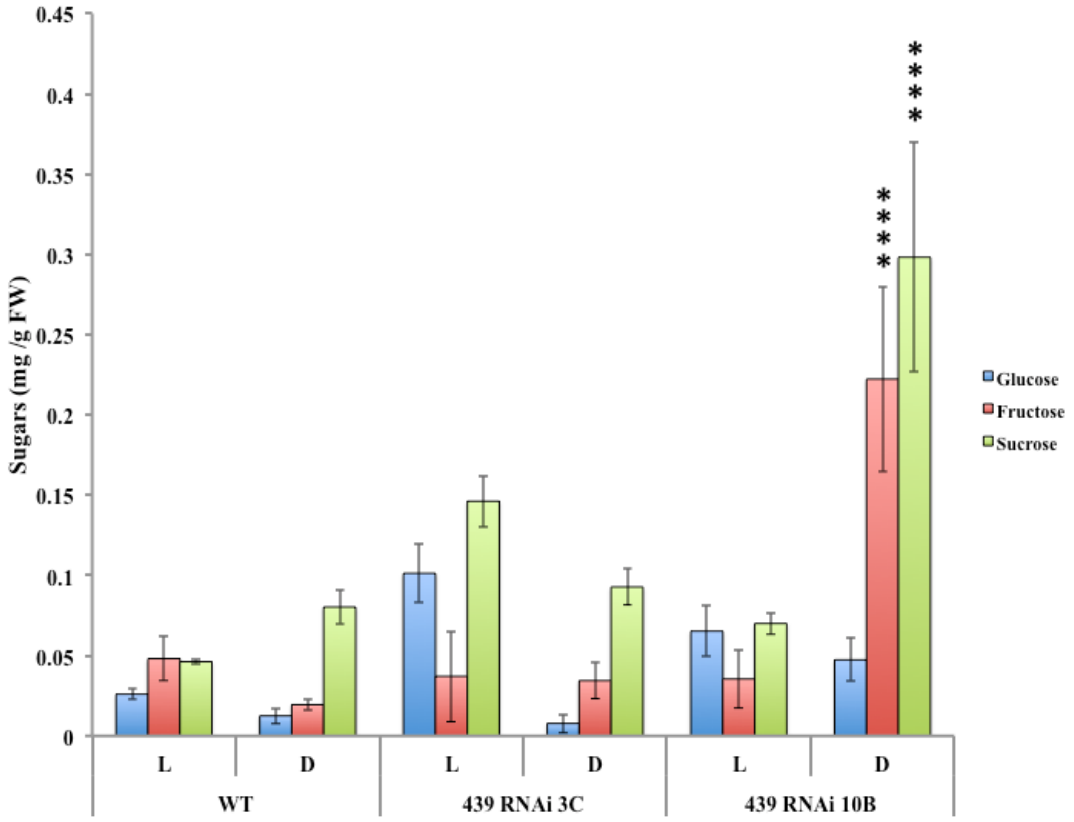

Figure 4.22. Levels of Glucose, Fructose and Sucrose in *KfMYB439* knockdown lines, at the end of the light (L; 23:00) and dark (D; 11:00) periods. *KfMYB439_RNAi_3C* shows no significant differences in sugar concentration when compared to wild type, whilst *KfMYB439_RNAi_10B* shows large increases in the level of fructose and sucrose during the dark period. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

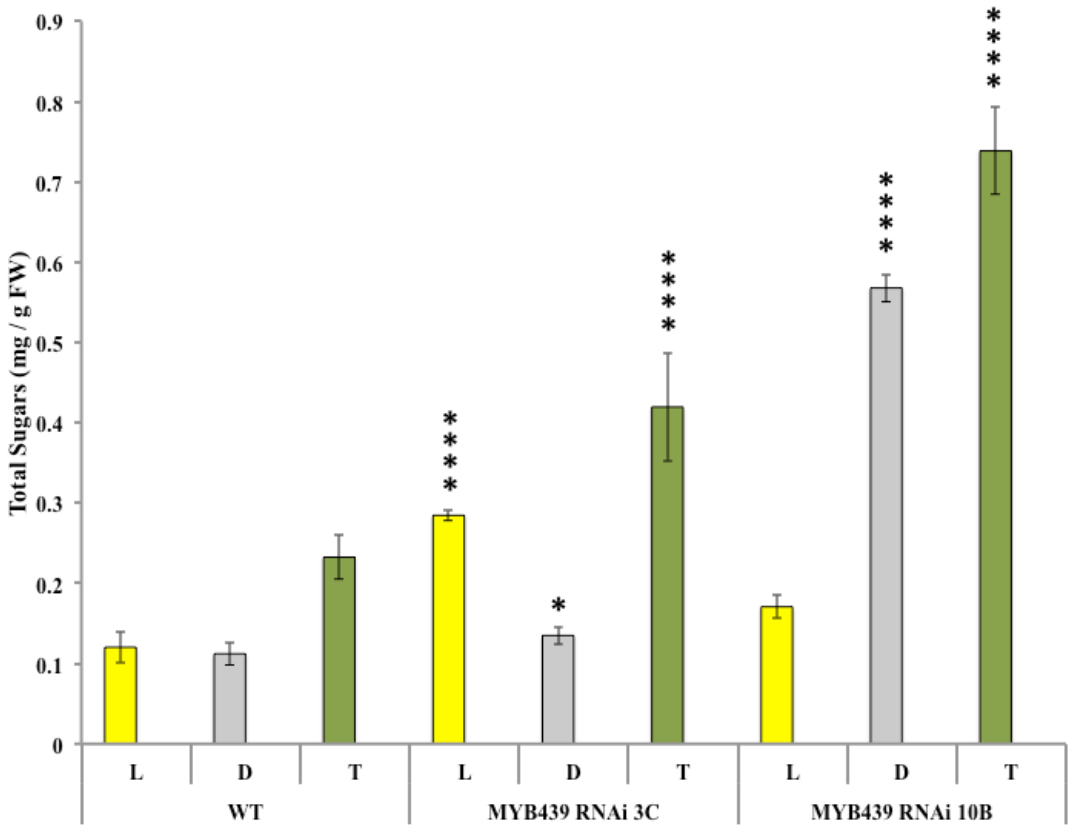

Figure 4.23. Variation in total soluble sugar levels in WT and the *KfMYB439 RNAi* lines. Yellow bars indicate light sugars levels (L), grey represent dark levels (D) and green bars show the total of sugar levels (T) calculated by summing the end-of-dark and end-of-light values for all sugars. There is a clear trend that these MYB439 RNAi lines have higher sugar levels during the day. *KfMYB439 RNAi 3C* shows lower sugar levels during the dark, whilst RNAi_10B has increased amount. Total sugar turnover in a day is higher than wild type. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

4.2.6 Immunoblot determination of the relative protein abundance of key enzymes associated with CAM

4.2.6.1 Protein abundance in *KfMYB439* full-length overexpression lines

In the light of the small changes in the transcript levels of only a few CAM-associated genes (Figs. 4.6 & 4.7), and the reductions in dark CO_2 fixation and Δ-malate in both over-expression lines (Fig. 4.16), it was important to determine whether the reductions in dark CO_2 fixation and Δ-malate were associated with decreases in the protein abundance of key CAM enzymes. Immunoblotting was used to examine the protein abundance of PEPC, phospho-PEPC, NAD-ME α and NAD-ME β, all measured using samples collected every 4 h over a light/ dark cycle as monitored for the associated transcript levels. Both *KfMYB439* over-expression lines had very similar protein levels to the WT for PEPC, NAD-ME α and NAD-ME β (Fig. 4.24A, C and D). Furthermore, the phosphorylation state of PEPC in the dark period, a direct result of the dark period increase in *PPCK* transcript levels and activity, and a key point of temporal regulation within the CAM pathway, was indistinguishable between the WT and the two FL over-expresser lines. These highly similar abundances for key CAM enzymes, and phosphorylation levels of PEPC in the dark, were consistent with the very similar levels of the corresponding transcripts detected using RT-PCR (Section 4.2.2.1). It is noteworthy that the small increase in *KfPPCK1* transcript levels, and their earlier rise in the dark period, compared to the WT (Figs. 4.6 & 4.7), was not associated with a phase advance in the detection of phospho-PEPC to the earlier part time points in the dark (Fig. 4.24B). This reveals that the small changes in transcript level were not sufficient to mediate changes in the nocturnal phosphorylation of PEPC; at least not within the detection limits of the immunoblotting.

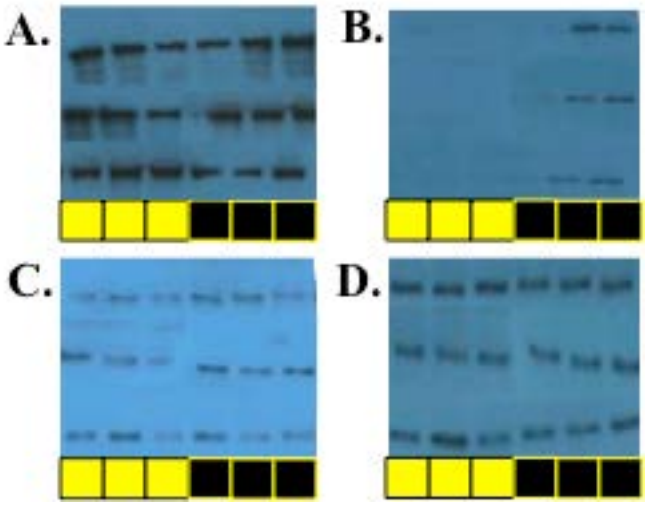

Figure 4.24. Immunoblotting revealed that several key CAM proteins did not change in abundance or phosphorylation state in the *KfMYB439_FL* over-expressor lines. **A.** Immunoblot for PEPC **B.** Immunoblot for phosphorylated PEPC; **C.** Immunoblot for NADME alpha **D.** Immunoblot for NADME beta. (Top –WT, Middle – 37C, Bottom – 54B).

4.2.6.2 Protein abundance in *KfMYB439* RNAi Knockdown lines

Immunoblot analysis of the *KfMYB439_RNAi* lines revealed clear differences in both the abundance of NAD-ME α, and the phosphorylation state of PEPC (Fig.**G** 4.25B and C). Specifically looking at phosphorylated PEPC, both RNAi lines showed lower levels of PEPC phosphorylation compared to the WT at the end of the dark period (22:00; 2 h before dawn), with line *RNAi_10B* lacking any detectable phospho-PEPC at 22:00 (Fig. 4.25B). This difference in phospho-PEPC in Figure 4.24B becomes even more striking when the amount of phospho-PEPC is compared to the total level of PEPC protein present in each sample (Fig. 2.25A). These results suggest these *KfMYB439_RNAi* lines have a reduced ability to achieve efficient phosphorylation of PEPC in the dark period compared to the WT. As transcript levels of *KfPPCK1* rose earlier in the dark period and were higher than WT at both 14:00 (2 h into dark) and 18:00 (middle of dark) in the RNAi lines, the reduced level and duration of dark phosphorylation of PEPC in the RNAi lines suggest that these lines have a problem

either with producing active PPCK from the transcript, or maintaining the PEPC in the phosphorylated state. This result could suggest that PEPC was being de-phosphorylated at a much more rapid rate than wild type in the RNAi lines (Section 4.2.2.2). Alternatively, PPCK is known to be inhibited by malate, and thus it could be that higher malate levels were inhibiting PPCK activity in the RNAi lines. Malate levels were markedly higher in the *RNAi_10B* line at the end of the light period (10:00, 2 h before dusk, Fig. 4.20). Assuming that malate level was sustained into the dark period, it may have been sufficient to inhibit PPCK. Similar findings have been reported previously for the *rNAD_ME* RNAi lines of *K. fedtschenkoi,* which also displayed reduced dark period phosphorylation of PEPC, but had similar levels of *KfPPCK1* transcript to the WT (Dever *et al.,* 2015). The *rNAD_ME* RNAi lines had higher levels of malate at the end of the light period due to their reduced ability to decarboxylate malate in the light. The authors therefore hypothesised that the elevated malate, relative to the WT, at the start of the dark period may have been sufficient to inhibit PPCK activity through the known ability of malate to inhibit the activity of this protein kinase (Dever *et al*., 2015). It is noteworthy that the largest decrease in dark phosphorylation in PEPC was observed for line *RNAi_10B* (Fig. 4.25B), which was the only *KfMYB439_RNAi* line to display a statistically significant elevation in the level of malate at the end of the light period (Fig. 4.20).

A difference in protein abundance was also detected for NAD-ME α. *KfMYB439_RNAi_3C* had a lower level of the protein throughout the 24h cycle, whilst *KfMYB439_RNAi_10B* only displayed reduced levels of this protein during the light period when compared to the wild type. These results are quite different to the transcript data obtained, suggesting that the change is due to alterations in post-

transcriptional modification, translation, protein turnover or post-translational changes (Section 4.2.2.2).

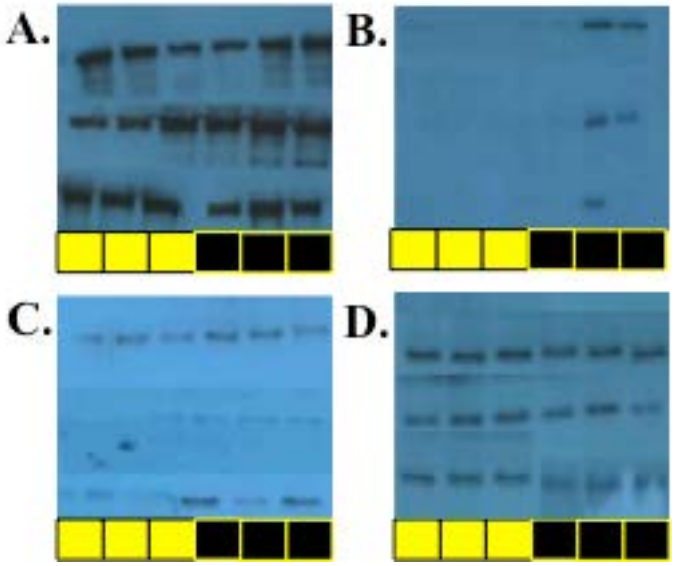

Figure 4.25. Western Blots used to identify protein level changes in key CAM enzymes due to reduction of *KfMYB439* transcript level. **A.** PEPC **B.** Phosphorylated PEPC; **C.** NADME alpha **D.** NADME beta. (Top –WT, Middle – 3C, Bottom – 10B).

4.2.7 Enzyme Assay conducted on transgenic *KfMYB439* lines

Whilst the immunoblots provided information about the abundance of several key CAM proteins, it was also important to investigate whether the perturbations of *KfMYB439* levels in the transgenic lines had an impact on the activity of CAM enzymes, especially in the light of the alterations in PEPC phosphorylation in the RNAi lines. The activity and malate sensitivity of PEPC was therefore measured in the WT and transgenic lines (Figs. 4.26 and 4.27 and Tables 4.6 and 4.7).

4.2.7.1 Enzyme assay results for *KfMYB439* full-length over-expresser lines

4.2.7.1.1 Total extractable activity of PEPC

To quantify PEPC enzyme activity, results were expressed relative to both total chlorophyll and total protein within the sample. Both were calculated because there is evidence that in transgenic plants levels of both chlorophyll and total protein can change, and so this enabled the enzyme activity to be compared using two different ways of standardisation.

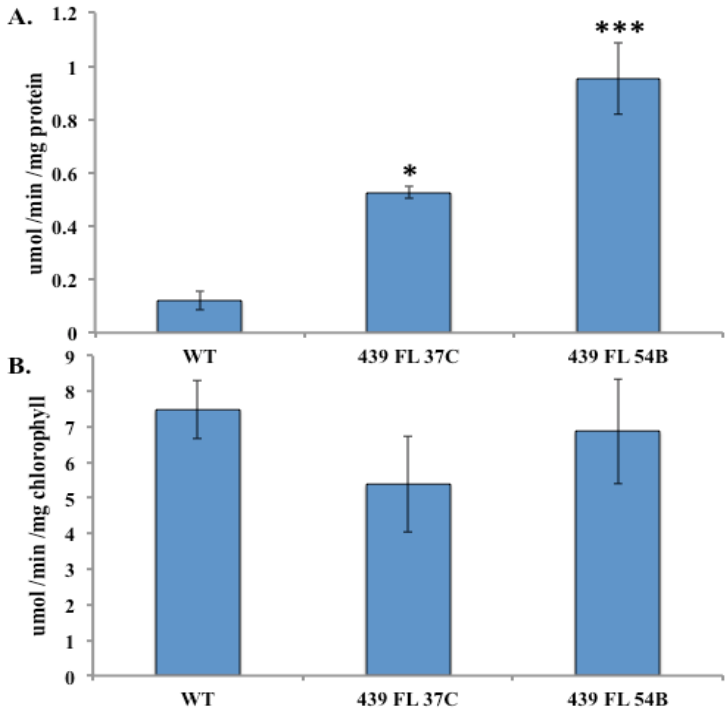

Figure 4.26. Total extractable activity of PEPC during the dark period (6h dark) for WT and 2 *KfMYB439_FL* lines. **A**. Total extractable activity of PEPC per mg of protein. **B.** Total extractable acitivity of PEPC per mg of Chlorophyll Both *KfMYB439* lines show vastly increased levels in activity compared to the wild type when activity is expressed per mg protein, but small average decreases in activity are seen when expressed per mg chlorophyll. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars above one column represent significant differences to the corresponding WT value.

For both *KfMYB439_FL* lines, the total extractable activity of PEPC per mg of protein was significantly higher than the WT activity (Fig. 4.26A; One-way ANOVA; 37C: P = 0.026, 54B: P = 0.0008). *KfMYB439_FL_37C* showed a 4.3-fold increase, whereas *KfMYB439_FL_54B* showed an increase of 7.8-fold. There was also a significant difference between the two *KfMYB439* lines, indicating that the higher the level of over-expression, the larger the increase in PEPC activity (One-way ANOVA, P = 0.021).

When comparing the total extractable activity of PEPC per mg of chlorophyll though, we now see that both *KfMYB439_FL* lines are not significantly different to wild type (Fig. 4.26B). From looking at average activity values per mg chlorophyll, there actually appears to be a small reduction in PEPC activity. These results are supported by more data collected from other experiments, such as transcript and immunoblot data. Therefore this may suggest that chlorophyll may be the more reliable measurement out of the two, and could suggest that total protein level is lower for some reason in transgenic lines.

4.2.7.1.2 Measurement of the malate sensitivity of PEPC

The malate sensitivity of PEPC depends on its phosphorylation state and is thus a key outcome of PEPC phosphorylation by PPCK in the dark. When PEPC is phosphorylated by PEPC in the dark period it becomes less sensitive to feedback inhibition by malate. The apparent K_i of extracted PEPC for malate was therefore determined. Consistent with published results, PEPC had a high K_i for malate in the dark, and a low K_i for malate in the light (Table 4.6). Both FL over-expressor lines possessed PEPC that had a lower K_i for malate than the WT in both the light and the dark (Table 4.6). In particular, the K_i for malate of the PEPC extracted from

KfMYB439_FL_54B was almost identical in the light and dark. Even for line KfMYB439_FL_37C the Ki only rose from 0.8 mM malate in the light to 1.4 mM in the dark, whereas the K_i of PEPC for malate in the WT rose from 1.9 mM in the light to 3.7 mM in the dark (Table 4.6).

Table 4.6. Variation in the apparent K_i of PEPC for malate in rapidly desalted extracts of CAM leaves of WT and the *KFMYB439_FL* transgenic lines sampled at the middle of the light (6h L) and the middle of the dark (6h D).

Line	Light Malate K_i (mM)	Dark Malate K_i (mM)
WT	1.9	3.7
MYB439 FL 37C	0.8	1.4
MYB439 FL 54B	0.8	0.9

4.2.7.2 Enzyme assay results for *KfMYB439* RNAi Knockdown lines.

4.2.7.2.1 Total extractable activity of PEPC

For both *KfMYB439* RNAi lines PEPC activity was also expressed revlative to both protein and chlorophyll. For both *KfMYB439_FL* lines the total extractable activity of PEPC per mg protein was significantly higher than the wild type (Fig. 4.27; One-way ANOVA; 3C: P = 0.025, 10B: P = 0.004). *KfMYB439_RNAi_3C* showed a 6.3-fold increase, whereas in line 10B PEPC activity increased 10-fold (Fig. 4.27).

When activity of PEPC is quantified per mg of chlorophyll we see that *KfMYB439_RNAi_3C* is now no longer significantly different to wild type, although from averages there is a slight increase (1.3 fold). On the other hand *KfMYB439_RNAi_10B* still has a significantly increased activity with a 1.8 fold increase (P = 0.007).

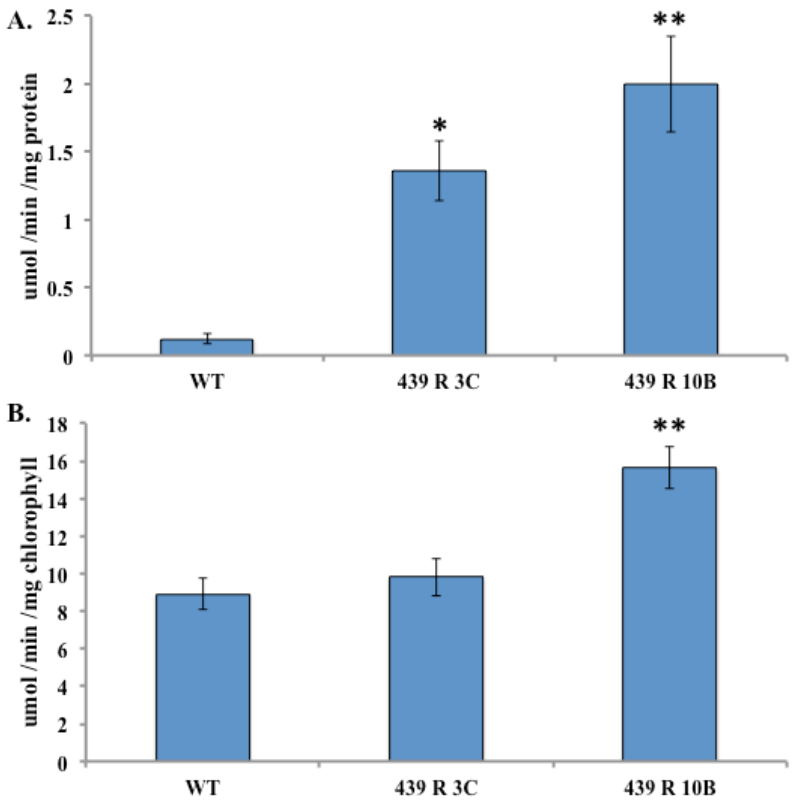

Figure 4.27. Activity level of PEPC during the dark period for WT and 2 *KfMYB439_RNAi* lines**, A.** Activity expressed per mg protein. **B.** Activity expressed relative to chlorophyll content. Both KfMYB439 lines show vastly increased levels in activity compared to the wild type relative to protein, but relative to chlorophyll, only *KfMYB439_RNAi_10B* is shown to have increased activity. Two-way ANOVA Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars above one column represent significant differences to the corresponding WT value.

4.2.7.2.2 Measurement of the Malate Sensitivity of PEPC

Line *KfMYB439_RNAi_3C* had a similar apparent K_i to wild type during the light period, whereas line *10B* had a much lower K_i than the WT, which was the same as both FL lines (Table 4.7). Both RNAi lines showed an increase in the K_i of PEPC for malate in the dark period, although whilst the WT reached a dark K_i of 3.7 mM malate, line 3C reached only 1.6 mM and line 10B reached only 1.5 mM (Table 4.7). Thus, the apparent K_i of PEPC for malate in the dark in both RNAi lines failed to reach even 50 % of the WT K_i. These findings indicate that PEPC in the RNAi lines

would be inhibited by malate much earlier during the dark period than observed in the WT, and this was consistent with the LD gas exchange data which did demonstrate the decline in nocturnal CO_2 fixation did commence earlier in both RNAi lines relative to the WT. These K_i results are also consistent with the phospho-PEPC immunoblot results, which indicated reduced dark period phosphorylation of PEPC in both RNAi lines (Section 4.2.4.2)

Table 4.7. Malate K_i values showing differences in levels of malate sensitivity between wild type and *KfMYB439_RNAi* lines.

Line	Light Malate K_i (mM)	Dark Malate K_i (mM)
WT	1.9	3.7
MYB439 RNAi 3C	1.3	1.6
MYB439 RNAi 10B	0.8	1.5

4.2.8 The impact of drought stress on the growth and stress-related physiology of the transgenic lines

Due to the importance of CAM to whole plant water use efficiency and the positive response of the plant to drought conditions, it was important to investigate the growth performance and physiological responses of the transgenic lines to drought stress treatment. Large clonally replicated populations of both WT and the transgenic lines were raised from developmentally synchronised leaf margin plantlets under greenhouse conditions. Plants were grown for 24 weeks well watered and then plant were either maintained in well-watered conditions for the remainder of the experiment or were subjected to drought stress by withholding watering. Plants were allowed to grow for a further 31 days and then sampled for the measurement of above ground wet and dry biomass yield, leaf anthocyanin content, leaf chlorophyll and carotenoid content and dawn and dusk malate levels.

4.2.8.1 Drought effects of *KfMYB439* full-length over-expresser lines

4.2.8.1.1 Above-ground biomass

For the biomass data, eight individual clonal plants were used in each growth condition for the WT whilst 6 plants were used for line *KFMYB439_FL_37C* and 9 plants were used for *KfMYB439_FL_54B*. These different numbers of biological replicates depended on the establishment and uniform growth of the leaf margin plantlets at the beginning of this growth and drought experiment. Some lines established better than other from plantlets meaning that more replicates were available for those lines.

For all lines, there was a significant difference between both fresh and dry weights when comparing each line in well watered and drought conditions (well-watered: WT

= 371g, 37C = 260g, 54B = 280g; Drought: WT = 76g, 37C = 62g, 54B = 68g; $P < 0.0001$; Fig. 4.28). When comparing between the WT and the transgenic FL lines, it can be seen that the above ground fresh and dry weight was significantly lower for both *KfMYB439_FL_37C* and *KfMYB439_FL_54B* under well-watered conditions (Fig. 4.28). WT dry weight yield was larger than the over-expresser lines when well-watered, but this difference was not observed when the lines had been subjected to drought-stress by withholding water for 31 days (Fig. 4.28; well-watered: WT = 18.1g, 37C = 12.6g, 54B = 12.7g, Drought: WT = 10.0g, 37C = 8.2g, 54B = 8.8g ; Fig. 4.28). There was no significant difference between dry weights for the two *KfMYB439* FL lines in well watered and drought-conditions, implying these lines performed similarly well in both conditions. From the percentage moisture calculations, all lines were found to have approximately similar levels of moisture in each watering condition (Well watered = 95%, Drought = 87%; Fig. 4.28E).

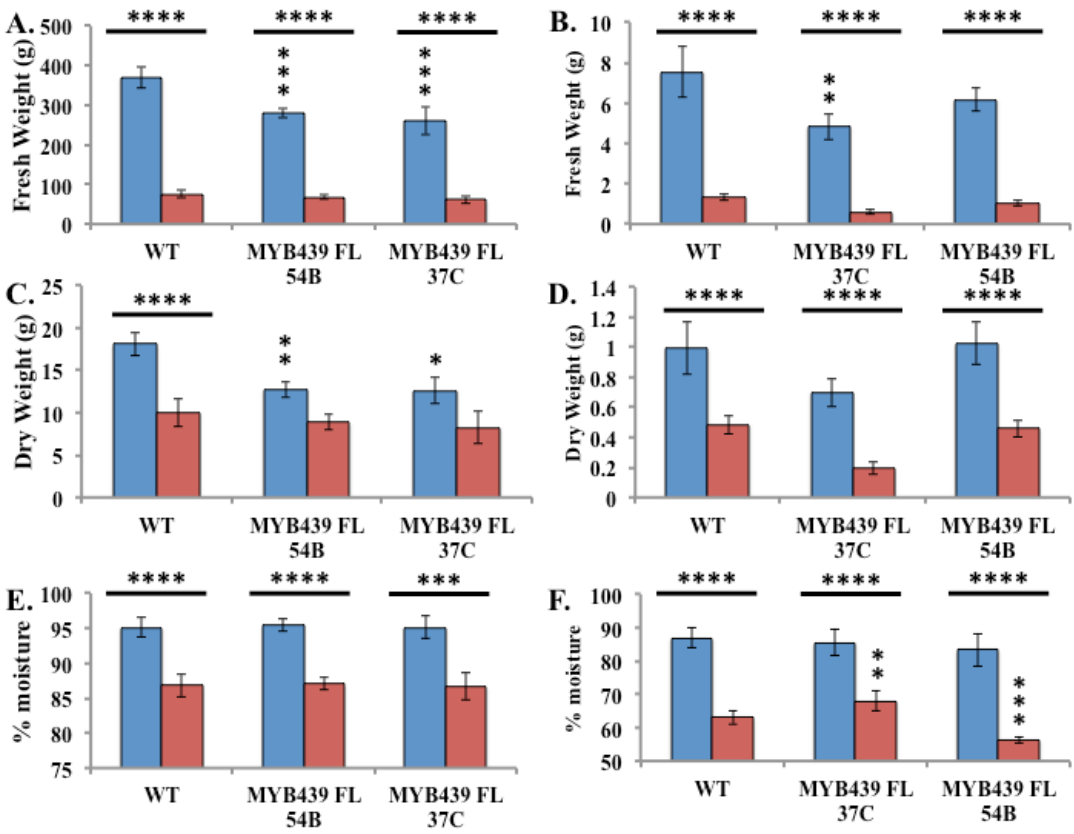

Figure 4.28. Fresh and dry weight and relative moisture content for well-watered and drought-stressed WT and FL lines. Blue bars show well watered lines, and red bars show drought-stressed lines. A, C & E show data collected for above ground weight, and B, D & F show data for below ground tissues (roots). A and B show fresh weight of samples at time of harvest, C and D, show dry weight after tissues were placed in a drying oven at 70˚C until a constant weight was reached. E and F show the % moisture held within the plant; calculated from the fresh and dry weights. A Two-Way ANOVA Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one column, e.g. well-watered/ drought represent significant differences compared to the corresponding WT value.

4.2.8.1.2. Below-ground root biomass

The roots of each plant were also collected and weighed. There were again significant differences in root fresh weight between well-watered and drought-stressed conditions (well-watered: WT = 7.5 g, 37C = 4.8 g, 54B = 6.2 g, Drought: WT = 1.3, 37C = 0.6g, 54B = 1.1g ; $P < 0.0001$: Fig 4.27B & D). Furthermore, the root biomass of line *KfMYB439_FL_37C* also showed a significant difference from the wild type in well-

watered conditions (Fig. 4.28B; P = 0.0064). However, the dry weight of the roots did not differ significantly between WT and *KfMYB439_FL* lines (well-watered: WT = 0.99 g, 37C = 0.71 g, 54B = 1.03 g; Drought: WT = 0.49 g, 37C = 0.19 g, 54B = 0.46 g). There were no significant differences in the moisture content of roots under well-watered conditions, but at the end of the drought-stress treatment there were significant differences (Fig. 4.28F). However, moisture content of the soil was very similar for each line (Appendix: Fig S4.1 & Table S4.1). Line *KfMYB439_FL_37C* roots had higher moisture content than the WT at the end of the drought treatment (P = 0.002), whilst *KfMYB439_FL_54B*, had significantly lower moisture content in its roots than the WT following the drought-stress treatment (Fig.4.28F; P = 0.0002).

4.2.8.1.3 Induction of UV protective anthocyanin pigmentation in response to drought-stress

Anthocyanin synthesis in response to abiotic stress is known to be regulated via the circadian clock in Arabidopsis. As *KfMYB439* was discovered as a CAM-induced and clock-controlled transcription factor, and was a member of the *RVE* family of circadian clock genes, it was important to investigate the anthocyanin levels of leaves from well-watered and drought-stressed plants. *K. fedtschenkoi* plants turn purple during drought-stress treatment due to high levels of anthocyanin pigments accumulating in the leaves (Fig. 4.29 A – D; Dever *et al.,* 2015). Total leaf anthocyanins were therefore measured in the WT and transgenic lines for the well-watered and drought-stressed plants grown in this experiment. Drought-stress caused a large and significant induction of leaf anthocyanin levels for all lines (P < 0.0001, Fig. 4.29). Both the WT and the FL lines possessed similarly low levels of leaf anthocyanins under well-watered conditions (WT = 0.001µmol cm^{-2}, 37C =

0.001µmol cm-^{2}, 54B = 0.003 µmol cm^{-2}). Under drought-stress, the FL lines accumulated slightly more anthocyanins than the WT, but this difference was only significant for *KfMYB439_FL_37C* (WT = 0.008 µmol cm^{-2}, 37C = 0.010 µmol cm^{-2}, 54B = 0.009 µmol cm^{-2}; 37C; P = 0.034).

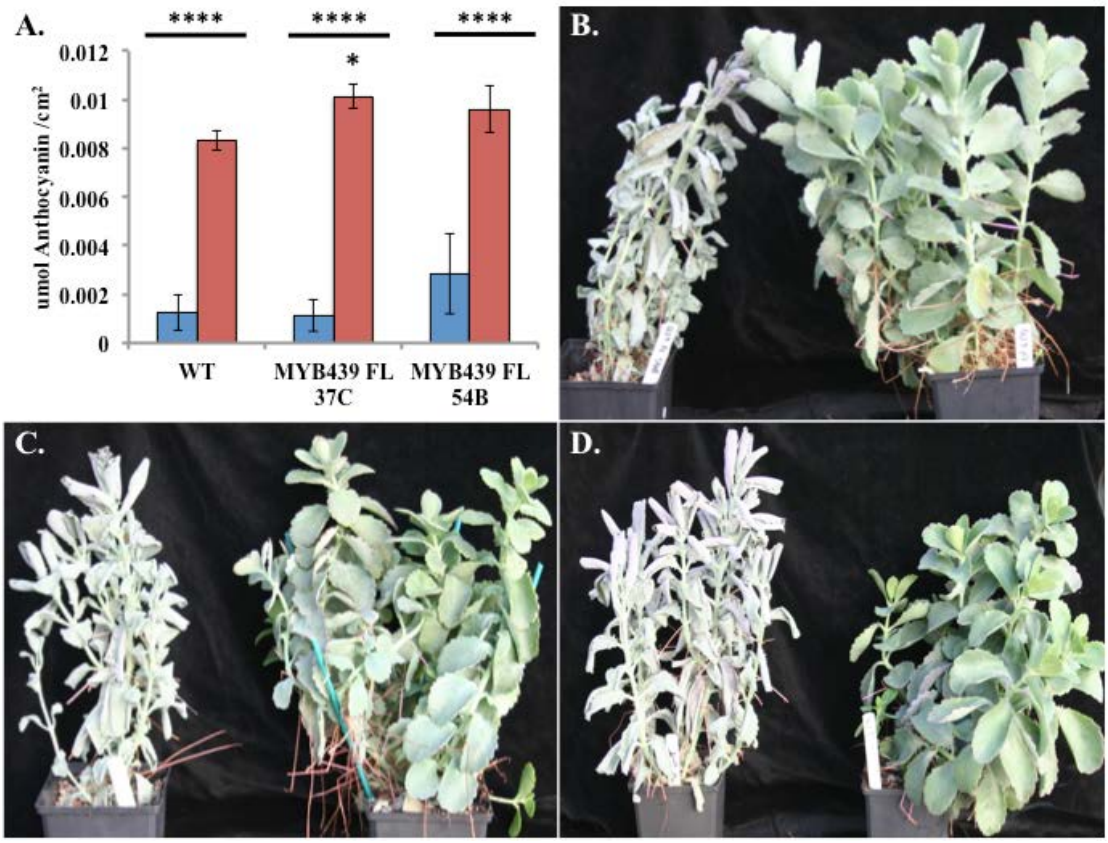

Figure 4.29. Variation in the level of anthocyanin leaf pigments in response to drought-stress for WT and *KfMYB439_FL_37C* and *54B*. A. Quantified levels of Anthocyanins in WT and the KfMYB439 FL lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one column, e.g. drought, represent significant differences to the corresponding WT value. B. Images of drought-stressed (Left) and well-watered (Right) WT. C. Images of drought-stressed (Left) and well-watered (Right) *KfMYB439_FL_37C*. D. Images of drought-stressed (Left) and well-watered (Right) *KfMYB439_FL_54B*.

In the light of the anthocyanin results, other leaf pigments were also measured for the well-watered and drought-stressed plants, including chlorophyll A, chlorophyll B and carotenoids (Fig. 4.30A-C). All three pigments increased in abundance with drought-stress, even after leaf area changes associated with drought stress were taken into account (Fig. 4.30; $P < 0.0001$). There were, however, no significant differences in the levels of these pigments between the WT and FL lines under either well-watered or drought-stressed conditions.

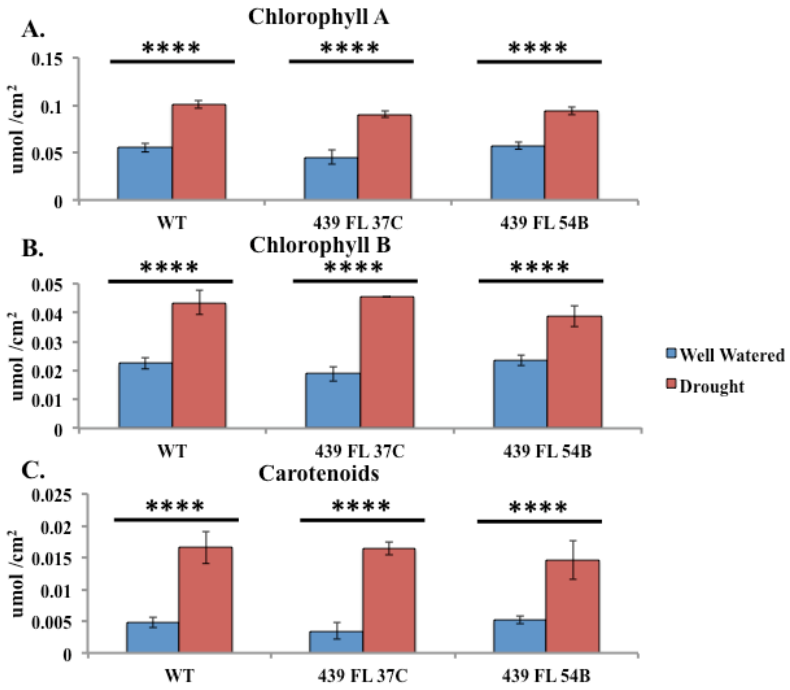

Figure 4.30. Variation in the levels of photosynthetic pigment in response to drought-stress in WT and *KfMYB439* FL lines. **A.** Changes in the levels of Chlorophyll A in response to drought. **B.** Changes in the levels of Chlorophyll B in response to drought. **C.** Changes in the levels of Carotenoids in response to drought. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; **** shows $P = \leq 0.0001$. Stars over bars at the top of the graph represent lines that are significantly different between treatments.

4.2.6.1.4 Metabolite levels in drought conditions

It was also important to investigate the impact of drought-stress on the magnitude of the daily CAM cycle. Levels of malate at dawn and dusk were measured in both the well-watered and drought-stressed lines using CAM leaves (LP6).

In well-watered conditions, all three lines accumulated similar amounts of malate at the end of the dark and had similar amounts of malate remaining at the end of the light (Fig. 4.31). The *KfMYB439_FL* lines had slightly more malate at the end of the light period than the WT (Fig. 4.31A). Following drought-stress, WT accumulated similar amounts of malate as it did when well watered, but both the FL over-expresser lines showed significant reductions in malate concentration compared to WT (Fig. 4.31B & D). The reduced level of malate accumulated by the *KfMYB439_FL* lines may be a sign that they were CAM-idling, whereby their stomata would be closed throughout the 24 h cycle in order to conserve leaf water content, and respired CO_2 would be recycled during the dark period yielding the reduced level of malate measured in the end of dark samples. However, this could only be confirmed by performing detailed gas exchange analysis on these lines as they progress from well watered to strongly drought-stressed plants. It was not possible to make these measurements on these plants due to time constraints and lack of availability of the IRGA.

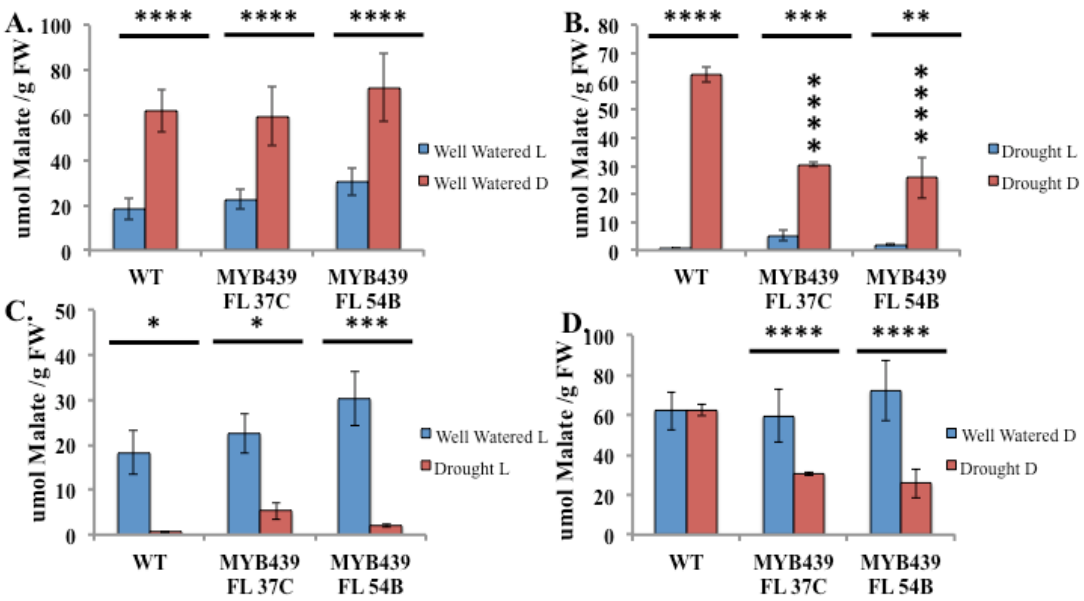

Figure 4.31. Daily fluctuations in malate content of CAM leaves of the WT and FL over-expressor lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. well-watered and Drought, and stars above one column, e.g. drought, represent significant differences from the corresponding WT value.

4.2.8.2 Drought effects of *KfMYB439* RNAi Knockdown lines

For the plant biomass measurements, as many biological replicates as possible were used in order to underpin more powerful statistical analysis. For the experiment to study the growth responses of the RNAi lines to drought-stress, eight WT plant were compared with seven plants of *KfMYB439_RNAi_3C*, and nine plant of line *RNAi_10B*. The minor variation in the number of biological replicates was unavoidable due to variation in the establishment and growth of the developmentally synchronised leaf margin plantlets used to set up the experiment.

4.2.8.2.1 Above-ground biomass

WT and *KfMYB439 RNAi* lines *3C* and *10B* were harvested after 31-days under either drought or well-watered conditions and samples collected for both above-ground and below-ground biomass measurements. Following drought-stress, the fresh weight measurements for both the WT and RNAi lines of *K. fedtschenkoi* revealed that drought-stress caused a significant reduction in fresh weight for all lines (well-watered: WT = 371 g, 3C = 321 g, 10B = 365 g; drought: WT = 76 g, 3C = 91 g, 10B = 108 g ; $P < 0.0001$, Fig 4.32.A). The fresh material was also dried in an oven at 70˚C until it reached a uniform dry weight. These dry weight measurements revealed that both of the *KfMYB439 RNAi* lines performed better under drought-stress than the WT (Fig. 4.32C). In particular, for line *KfMYB439_RNAi_10B* there was no significant difference in the dry weight of the above-ground tissues between the two growth conditions. (well-watered: WT = 18.1g, 3C = 17.2g, 10B = 15.8g, Drought: WT = 10.0g, 3C = 7.2g, 10B = 9.6g Fig 4.32.C). All lines had approximately the same percentage moisture in their above-ground tissues in each condition (Fig 4.32.E).

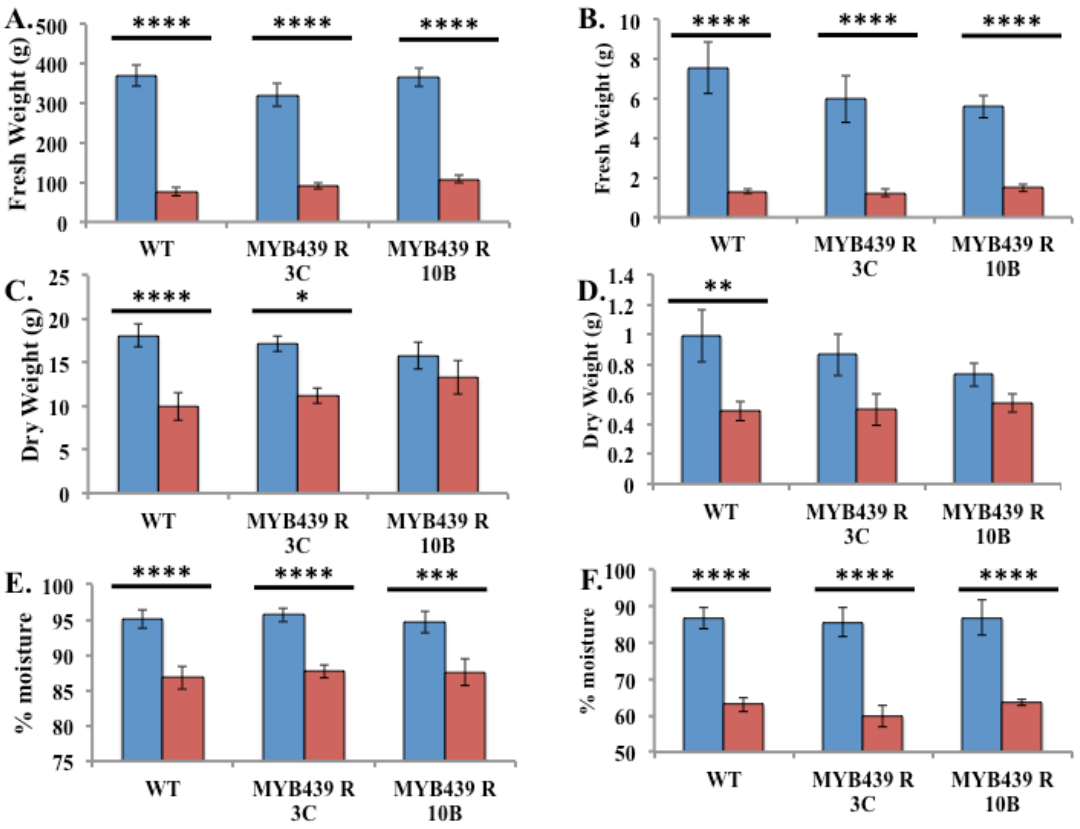

Figure 4.32. Fresh and dry weight and relative moisture content for well-watered and drought-stressed WT and RNAi lines. Blue bars show well watered lines, and red bars show drought-stressed lines. A, C & E show data collected for above ground weight, and B, D & F show data for below ground tissues (roots). A and B show fresh weight of samples at time of harvest, C and D, show dry weight after tissues were placed in a drying oven at 70˚C until a constant weight was reached. E and F show the % moisture held within the plant; calculated from the fresh and dry weights. A Two-Way ANOVA Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one column, e.g. well-watered/ drought represent significant differences compared to the corresponding WT value.

4.2.8.2.2. Below-ground root biomass

Roots were also removed and weighed for plant. For WT and both *KfMYB439_RNAi* lines, the root fresh weights at time of harvest were drastically lower in drought conditions than in well watered (Well watered: WT = 7.5 g, 3C = 6.0 g, 10B = 5.6 g, Drought: WT = 1.3 g, 3C = 1.2 g, 10B = 1.5 g $P < 0.0001$; Fig. 4.32B).

Looking at the root dry weights, *KfMYB439_RNAi_10B* did not show a significant difference between the two growth conditions, although the average dry weight was still lower following drought-stress for this line (well-watered: WT = 1.0 g, 3C = 0.86 g, 10B = 0.73 g, Drought: WT = 0.49 g, 3C = 0.50 g, 10B = 0.54 g).

KfMYB439_RNAi_3C also showed no significant difference in its root dry weight (Fig. 4.32D).

In terms of the percentage moisture of the roots, there was only a large significant difference between each line in the two growth conditions (i.e. roots of each line contained much less water following drought-stress), but not between the lines ($P < 0.0001$, Fig. 4.32F).

4.2.8.2.3 Induction of UV protective anthocyanin pigmentation in response to drought-stress

In well watered conditions wild type only has a small amount of anthocyanins present, due to the plant not being under stress (0.001 uM/cm^2), With the *KfMYB439_RNAi* lines, there appears to be very low amounts of anthocyanin present, even lower than WT (Both concentrations are < 0.0001uM/cm^2 , 3C: $P = 0.040$, 10B: $P = 0.041$; Fig. 4.33). Once the plants have been subjected to 31-days drought though, all three show vast increases in levels, and all three increase to approximately the same value. This also means that both *KfMYB439_RNAi* lines have produced more

anthocyanin, since drought induction (WT = 0.008uM/cm^2, 3C = 0.009uM/cm^2, 54B = 0.008 uM/cm^2).

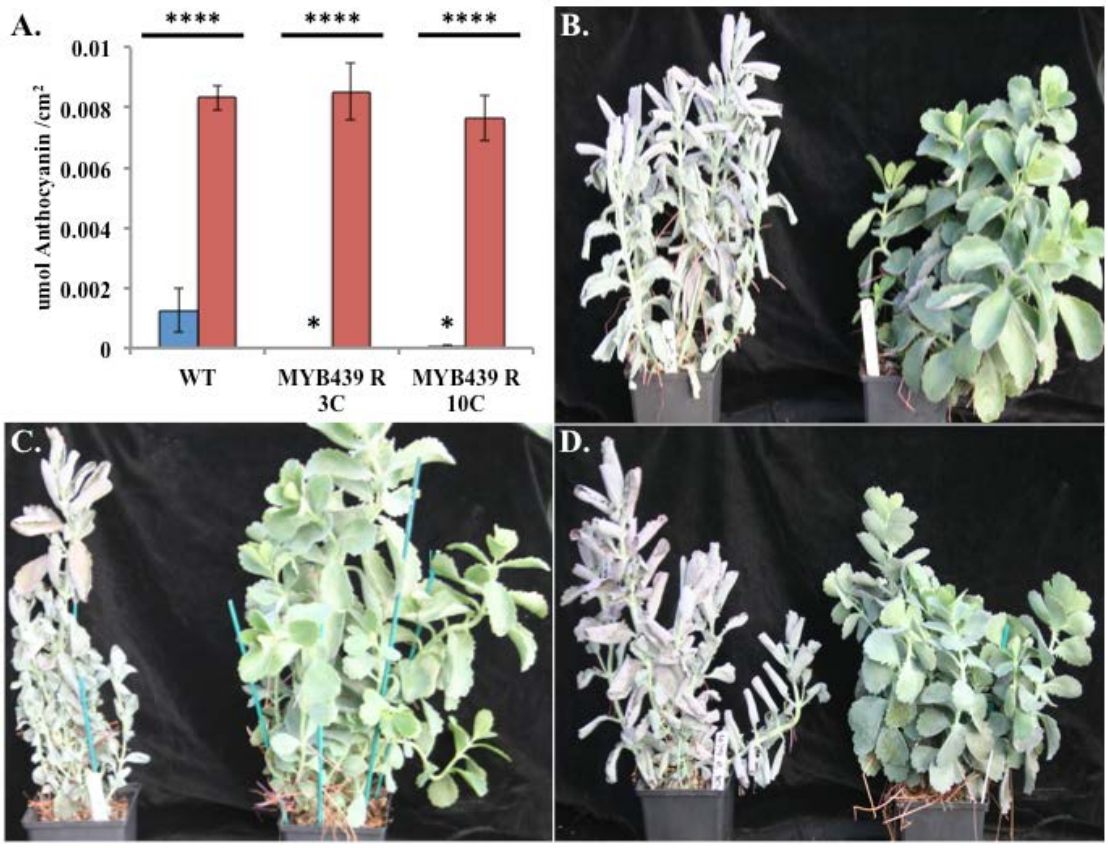

Figure 4.33. Variation in the level of anthocyanin leaf pigments in response to drought-stress for WT and *KfMYB439_RNAi_3C* and *10B*. **A.** Quantified levels of Anthocyanins in WT and the KfMYB439 RNAi lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one column, e.g. drought, represent significant differences to the corresponding WT value. **B.** Images of drought-stressed (Left) and well-watered (Right) WT. **C.** Images of drought-stressed (Left) and well-watered (Right) *KfMYB439_RNAi_3C*. **D.** Images of drought-stressed (Left) and well-watered (Right) *KfMYB439_RNAi_10B*.

As with the overexpressing lines, *KfMYB439_RNAi*s show a large increase in amounts of other pigments in the leaves; Chlorophyll A, B and Carotenoids ($P < 0.0001$; Fig. 4.34).

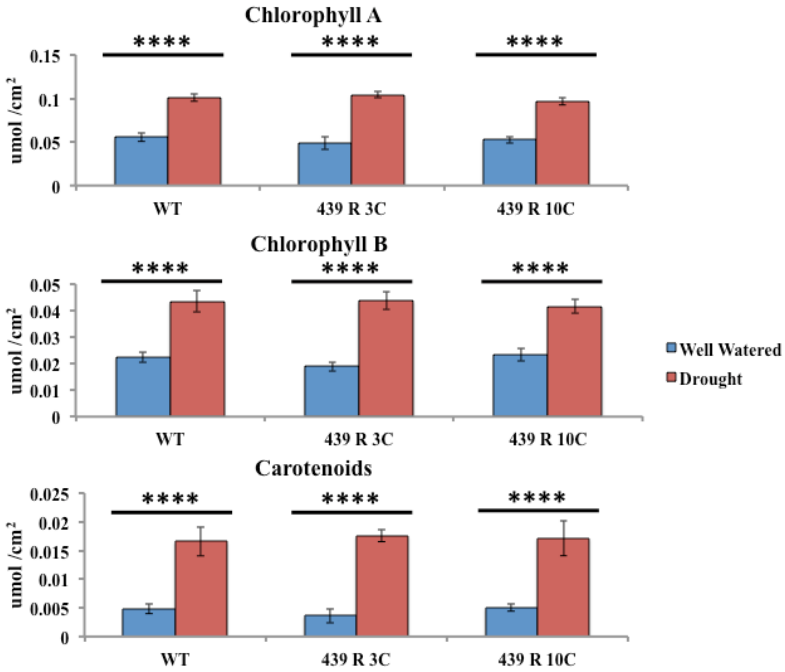

Figure 4.34. Variation in the levels of photosynthetic pigment in response to drought-stress in WT and *KfMYB439* RNAi lines. **A.** Changes in the levels of Chlorophyll A in response to drought. **B.** Changes in the levels of Chlorophyll B in response to drought. **C.** Changes in the levels of Carotenoids in response to drought. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; **** shows $P = \leq 0.0001$. Stars over bars at the top of the graph represent lines that are significantly different between treatments.

4.2.8.2.4 Metabolite levels in drought conditions

The level of malate was assayed for indications into the level of CAM occurring during well watered and droughted lines. For all lines there is significant differences between light and dark for each line in well watered, as expected if full CAM is occurring ($P < 0.0001$, Fig. 4.35A).

Once the lines are placed into drought, *KfMYB439_RNAi_10B* malate levels drop dramatically in both the light and dark, resulting in there no longer being a significant difference between the two. With *KfMYB439_RNAi_3C*, although there is still

significant differences between light and dark ($P < 0.0001$), the amount of malate accumulated has more than halved, whereas WT has stayed roughly the same (Fig. 4.35B & D). It therefore again appears that these lines have also started to conduct CAM idling.

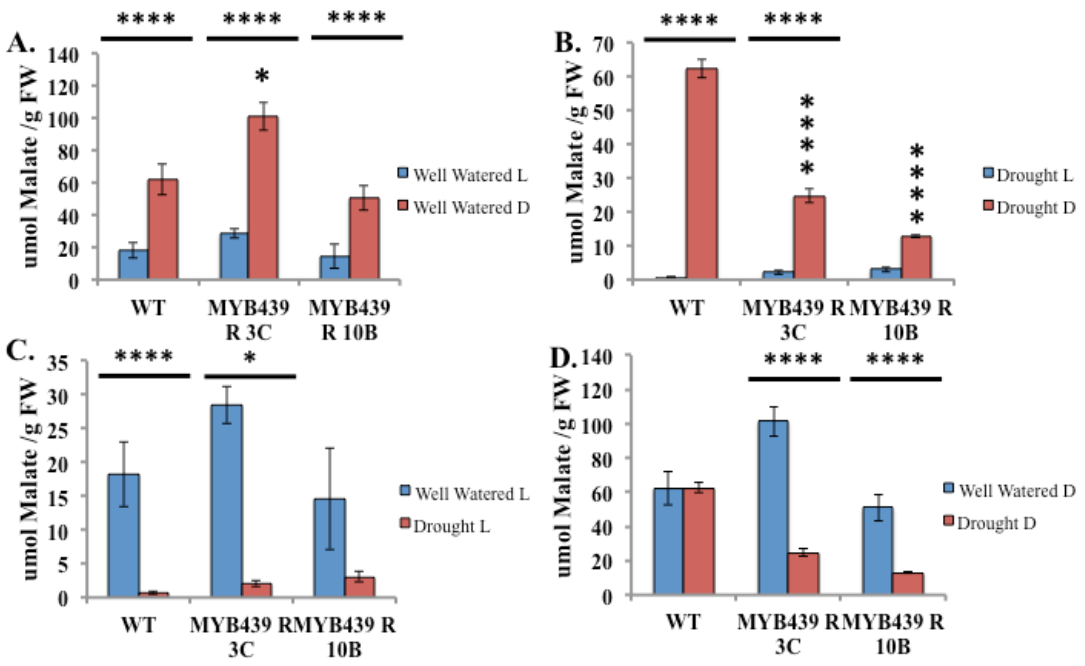

Figure 4.35. Daily fluctuations in malate content of CAM leaves of the WT and RNAi knockdown lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. well-watered and Drought, and stars above one column, e.g. drought, represent significant differences from the corresponding WT value.

4.2.9 *KfMYB439*'s impact on reproductive success

In order for *K. fedtschenkoi* to thrive in the semi-arid environment to which it is native in Madagascar, it must reproduce effectively. To study the impact of manipulating *KfMYB439 levels* on the reproductive success of the various FL and RNAi lines, leaves were removed from each *KfMYB439* transgenic line and were left in dry trays in the greenhouse to allow them to produce plantlets along the leaf margins. As each notch around the edge of a *K. fedtschenkoi* leaf is capable of producing a clonal plantlet, the ratio of notches/ plantlet was calculated to determine how many notches led to the production of 1 plantlet.

4.2.9.1 *KfMYB439* full-length overexpression lines impact on reproductive success

For the *KfMYB439_FL* lines, both had a lower average number of notches per plantlet formed than the WT (Fig. 4.36). This suggested that they were slightly more successful at reproducing than WT. *KfMYB439_FL_37C* displayed a large range of values, as indicated by the large standard error bar (Fig. 4.36). Thus, the difference between *FL_37C* and the WT was not significant. However, for line *KfMYB439_FL_54B*, there was a significant increase in the number of plantlets produced; a lower number of notches was required per plantlet produced (i.e. more notches generated plantlets on the margin of the *FL_54B* line compared to the WT; P = 0.0024).

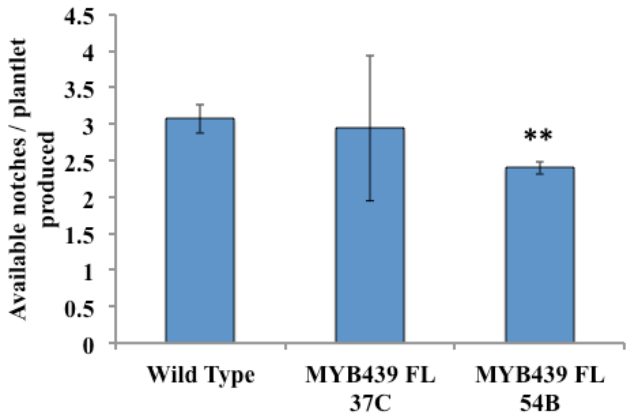

Figure 4.36. Impact of over-expression of *KfMYB439* on reproductive output in the form of leaf margin plantlet formation (For WT n = 39 leaves, 37C n = 18 leaves, for 54B n = 37 leaves). Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars above each column represent significant differences to the corresponding WT value.

4.2.9.2 *KfMYB439* RNAi knockdown lines impact on reproductive success.

For the *KfMYB439* RNAi lines, it was again found that fewer notches were required per plantlet formed, revealing that a larger reproductive effort was achieved by the detached leaves from these lines (Fig. 4.37). *KfMYB439_RNAi_3C* came very close to each leaf margin notch producing a plantlet (a ratio of 1.3), indicating it was capable of producing many more plantlets per leaf on average than the WT (P < 0.0001). *KfMYB439_RNAi_10B* also had a lower average number of notches per plantlet than the WT, although the difference was not significant for this line due to the large standard error (Fig. 4.37).

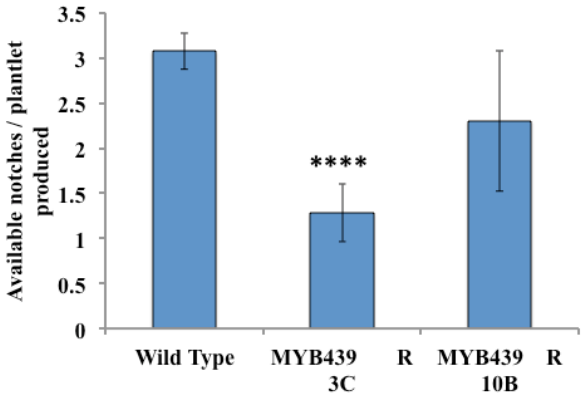

Figure 4.37. Impact of RNAi-based silencing of *KfMYB439* on reproductive output in the form of leaf margin plantlet formation (For WT n = 39 leaves, 3C n = 21 leaves, for 10B n = 18 leaves). Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars above each column represent significant differences to the corresponding WT value.

4.3 Discussion

This chapter describes the detailed investigation of role of *KfMYB439* in the circadian optimisation of CAM in *K. fedtschenkoi* through the use of transgenic lines in which *KfMYB439* was either silenced using a hairpin RNA RNAi silencing construct, or constitutively over-expressed using a 35S promoter driven full length open reading frame construct. Two of the best RNAi and over-expression lines were selected from the larger original populations of transgenic lines (Chapter 3). In particular, the two chosen lines were selected due to them displaying varying levels of over-expression and knockdown of the target *KfMYB439* gene, which allowed the impact of gene dosage to be investigated. However, there were not many *KfMYB439_RNAi* lines to choose from due to a high proportion of plantlet death during the regeneration of these RNAi transgenic lines. The four chosen lines; *KfMYB439_FL_ 37C* and *54B*, and *KfMYB439_RNAi_3C* and *10B*, were subjected to various molecular, biochemical and physiological phenotyping tests, to underpin the further elucidation of the function (if any) of this transcription factor in the daily cycle of CAM in *K. fedtschenkoi*.

4.3.1 Bioinformatic characteristation of *KfMYB439*

Identification of the family and sub-family to which a plant transcription factor belongs can be very useful for helping to define its potential functions. From initial analysis of the longest cDNA sequence available from assembled 454 sequence reads generated from a wide range of cDNAs from *K. fedtschenkoi*, *KfMYB439* was determined to be most similar to a gene in Arabidopsis called *HOMEODOMAIN-LIKE SUPERFAMILY PROTEIN/ REVEILLE 8* (*HSP*, *RVE8*, At3G09600), which is involved in regulating the circadian clock by modulating histone 3 (H3) acetylation

and regulates evening phased genes by binding to their EE (Farinas *et al.,* 2011; Hsu *et al.,* 2013). There is a large body of published work on the *REVEILLE* family of transcription factors in Arabidopsis, not least because the core circadian clock genes *LHY/CCA1* are members of the larger RVE family. It has been reasoned, and now shown, that LHY/CCA1 are not the only MYB-like TFs that regulate the expression of LHY and CCA1 target genes (with an evening element in their promoter) (Carre *et al.,* 2002; Hsu *et al.,* 2013). Therefore due to high sequence homology, other *RVE* transcription factors are of interest, as functional redundancy has already been demonstrated between some of the family members; for example *RVEs 4*, *6* and *8* (Hsu *et al.,* 2013).

Within the *RVE* family, a sub-family has been shown; namely the *LCL* sub-family; which have been shown to conduct different roles to other *RVE* members. From the bioinformatics analysis presented in this chapter, both the MYB domain, containing the SHAQKYF amino acid sequence, and the LCL domain were highly conserved between the *Arabidopsis RVE* family and *KfMYB439* (Fig. 4.2). Furthermore, from BLAST scores using the TAIR database (www.arabdidopsis.org), *KfMYB439* was shown to have the highest percentage similarity to *RVE8* in Arabidopsis, which has been shown to be involved in the circadian optimisation of clock processes, by working antagonistically with CCA1 (Farinas *et al.,* 2011). In Arabidopsis, RVE8 binds to the EE in the promoters of its target genes and activates transcription (Farinas *et al.,* 2011). By contrast, *CCA1/LHY* mainly repress the expression of the genes to which they bind at the EE such as binding to the *TOC1* promoter and therefore preventing acetylation of H3 to enable transcription (Perales *et al.,* 2007). There are also genes which have expression induced by CCA1/LHY, such as PRR 7&9 (Farre *et al.,* 2005).

For the analysis presented here from the SOLiD RNA-seq dataset for C_3 and CAM leaves of *K. fedtschenkoi* sampled every 4 h over a 12:12 LD cycle, all 9 *K. fedtschenkoi RVE* family members were detected as expressed in the dataset (Fig. 4.4). From Figure 4.4, the relative abundances of each *RVE* transcript found in the *K. fedtschenkoi* genome can be seen for both C_3 (A) and CAM (B) tissues. This demonstrated that many of the *RVE* family TFs in *K. fedtschenkoi* only increased in transcript abundance slightly with CAM induction. For example, for the *KfCCA1/LHY* genes, KF02345 showed a 1.8-fold induction, and KF106790 showed 2-fold induction. Other RVEs, including *KfMYB439* (*KF79015*), *KF37745* and *KF84360* displayed much greater fold-increases of 6.39, 4.48 and 9.67, respectively (Fig. 4.4 and Table 4.3). This therefore indicates that these genes are more likely to function in CAM circadian control in some way, as the *CCA1/LHY* genes that are known to function in the core *CCA1/LHY-TOC1* feedback loop of the circadian oscillator showed relatively modest changes in transcript abundance with CAM development in LP6 (Fig. 4.4). These data open up the possibility that these three *RVE* family genes (*KF79015*, *KF37745* and *KF84360*), that displayed large fold-inductions in CAM leaves relative to C_3, which includes *KfMYB439* studied in detail in this chapter, might either work redundantly or antagonistically. The potential regulatory interactions between this subset of strongly CAM-induced *RVE*s in *K. fedtschenkoi* certainly merits more detailed investigation in the future through experiments such as double and triple knockdown transgenic lines, and protein-protein interaction studies such as yeast two-hybrid analysis, and Fluorescence Resonance Energy Transfer imaging experiments. It would also be fascinating to study the promoter targets of each of these *RVE*s through Chromatin Immunoprecipitation followed by sequencing

(ChIP-seq) experiments. Lines for this purpose for *KfMYB439* have already been created.

The regulation of the *RVE* genes was also investigated using a recently obtained (July 2014) large Illumina RNA-seq light/dark timecourse dataset for CAM leaves, which compared *K. fedtschenkoi* wild type with 4 distinct transgenic lines that either had reduced levels of key CAM genes due to an RNAi transgene (*rPPCK1_3*, *rNAD_ME* and *rPPDK*), or over-expressed the iceplant central clock gene *TOC1* from the 35S promoter (*35S::McTOC1_OX*). The circadian rhythm of gas exchange in LL conditions becomes arrhythmic in transgenic lines *rPPCK1_3*, *rNAD_ME* and *TOC1_OX*, but remained rhythmic in line *rPPDK* (Boxall *et al*., 2015 in preparation; Dever *et al*., 2015; Dall'omo, 2011). Thus, this whole transcriptome RNA-seq dataset for these *K. fedtschenkoi* transgenic lines held the potential to provide novel insights into the response of *KfMYB439* and the other *RVE* family members in *K. fedtschenkoi* to genetic perturbations that influence the free-running rhythms of CAM. The *TOC1_OX* line was of particular interest as *TOC1* has been shown to feedback and inhibit *CCA1/LHY* expression in the plant circadian clock (Gendron *et al.,* 2012). This was also found to be the case for *K. fedtschenkoi*, as both *CCA1/LHY* genes showed constitutively lower levels of transcript compared to wild type in the *TOC1_OX* line (Fig. 4.5). In support of the idea that *RVE*s are able to function redundantly, it is also interesting that five out of the other seven *RVE*s, including *MYB439,* also showed a decrease in transcript abundance in the *TOC1_OX* line (Fig. 4.5). Other *RVE*s have been shown to play antagonistic roles in the Arabidopsis clock, and the data indicates that a candidate in *K. fedtschenkoi* is KF84360, as this was the only gene to increase in transcript abundance in the *TOC1_OX* line.

The effect of knocking down PPCK (in RNAi line *rPPCK1_3*) on the regulation of the *K. fedtschenkoi RVE* family genes was interesting (Fig. 4.5). The loss of clock-mediated PEPC phosphorylation in the dark in this line resulted in a decrease in the transcript level of most members of the *RVE* family, except for the two *CCA1/LHY* genes (KF02345 and KF106790), which showed a small increase. The *RVE* family may show a general decrease here because many MYB-related transcription factors, including RVEs, have been shown to play roles in metabolism and the clock, such as RVE2 and RVE7, which function within the clock regulatory feedback loops (Zhang *et al.,* 2007) and also RVEs 4, 6 and 8, which when all are mutated display severe circadian phenotypes (Hsu *et al.,* 2013). The *rPPCK1_3* line performs less dark CO_2 fixation than the WT due to the much more rapid inhibition of PEPC by malate as the night progresses. It may therefore be that feedback signal transduction pathways linking the clock to changes in CAM associated metabolism may be communicating key information about this lower level of dark CO_2 fixation to the core clock through the observed fall in the transcript abundance of many of the *RVE* family genes.
The majority of *RVE* family members showed a small decrease in the *rNAD_ME* RNAi knockdown line, except *KF130900*, which showed an increase in the light, although this was the *RVE* gene that showed the lowest amplitude of circadian regulation, and had the lowest transcript abundance (Fig. 4.5). The rNAD-ME line has a dramatic reduction in its ability to decarboxylate malate in the light period (Dever *et al.,* 2015). It is therefore assumed that this line struggles to maintain a high partial pressure of CO_2 behind closed stomata in the light, which in turn results in increased stomatal opening during the light period (Dever *et al.,* 2015). It is therefore possible that the *RVE* family may be involved in signalling to the clock, and so in overriding the circadian control of the stomata.

For the *rPPDK* line, there was a general trend of increased transcript for the *RVE* family, with only *KF73515* showing a decrease. Again, this line has been shown to perform little, if any, dark CO_2 fixation, and have a large reduction in dawn/ dusk malate oscillations, as a result of its reduced ability to convert pyruvate to PEP in the light during malate decarboxylation (Dever *et al.,* 2015). This line, like the *rNAD-ME* line, also fixed more CO_2 during the light period, and so the small transcript increases detected for the *RVE* genes may again be signalling via the circadian clock to enable stomatal opening in the light.

These data therefore support the original proposal that these *RVE* genes may be involved in feedback gene regulatory loops within the core clock, and in linking both the clock to CAM and perhaps also feeding back information from the metabolites associated with CAM into the core clock.

4.3.2 The impact of *KfMYB439* on gene expression

For both *KfMYB439* overexpressing lines, 37C and 54B, and both RNAi knockdown lines, 12A and 30A*, KfTOC1* transcript levels were not significantly different to the wild type. This is consistent with results for the orthologous RVE family TF in Arabidopsis, where loss of *RVE8* function did not significantly alter transcript level (Hsu *et al.,* 2013). *KfCCA1* also showed the same expression profile as seen in the WT, indicating that this TF may not play as close a role in the core clock as its *Arabidopsi*s ortholog*, RVE8.* In Arabidopsis, *RVE8*, as well as *CCA1/LHY,* has been shown to be able to inhibit its own gene expression (Wang *et al.,* 1998; Carre *et al.,* 2002). Thus, the excess levels of *KfMYB439* in the over-expressing lines are likely to cause inhibition of the production of the endogenous transcript. Furthermore, when considering the RNAi lines, it has been demonstrated in Arabidopsis that there is

genetic redundancy between *RVEs 4*, *6* and *8 (Hsu et al., 2013).* It therefore seems likely that in the *KfMYB439_RNAi* lines, *RVE4* and *RVE6* could be compensating for the reduction in *KfMYB439*.

As *KfCCA1* and *KfTOC1* transcript levels for the WT and *KfMYB439* transgenic lines showed little or no variation under LD cycles, even small variations in the levels and light/ dark temporal profile for other genes compared with the WT profile are of interest. All of the screened genes showed changes in their transcript levels at certain times of the day, suggesting that line *KfMYB439* may have a role in the clock, or in the clock's output. This could explain the widespread effect on various CAM genes. To further investigate the potential role of *KfMYB439* in the *K. fedtschenkoi* clock, it would be of interest to look at transcript abundance levels for the *PSEUDO RESPONSE REGULATOR* (*PRR*) genes. Hsu *et al*. (2013) predicted that *RVE*s and *PRR*s make up a new branch of the core clock, whereby *RVE4*, *6* and *8* work together to induce *PRR1* (*TOC1*), *5*, *7* and *9*, and in turn the *PRR* proteins repress the *RVE*s.

In constant free-running conditions of constant light, temperature and humidity (LL), circadian rhythms can be probed, and perturbations become more obvious. These conditions were therefore used to screen the *KfMYB439* transgenic lines for phenotypes.

In line *KfMYB439_FL_37C* there were constantly higher levels of *KfMYB439* transcripts throughout the 3-day LL timecourse. However, the temporal pattern of the transcript levels showed clear differences to the WT. Levels of *KfMYB439* transcripts peaked at 42 h LL, and, interestingly, at this time point both *KfCCA1* and *KfTOC1* transcript levels dropped below those in the WT. This result was particularly striking

for *KfCCA1*, which up to that time point had displayed a transcript profile very similar to the WT. Hsu *et al.*, demonstrated that *RVE8* levels in *Arabidopsis* are gated by the clock, and that the effects of over-expression of *RVE8* varied throughout the 24 h cycle (2013). Furthermore, in LL conditions it may take time for *KfMYB439* to overcome its repression before there is a high enough protein concentration to affect its target genes. This therefore supports the proposal that *KfMYB439* functions somewhere in the central circadian clock, like its ortholog in Arabidopsis does. The fact that the effects could only be seen during constant LL conditions, supports this proposal because it is only under LL free-running conditions that circadian efficiency can be probed effectively.

KfPPCK1 transcript levels were higher than the WT in the *KfMYB439_FL* lines under LD conditions. Under LL constant conditions, *KfPPCK1* displayed a dampened amplitude, suggesting that a constant high level of *KfMYB439* was in some way able to supress the periodic peaks of *KfPPCK1* accumulation under constant conditions. This could also be interpreted as evidence that the circadian clock was no longer functioning in a robust manner, leading to dampening of output rhythms such as that in *KfPPCK1* transcript levels.

When line *KfMYB439_RNAi_10B* was studied under LL free-running conditions, the RNAi knockdown of the target *KfMYB439* transcript was not detected. By contrast, by 42 h into the LL experimental run, *KfMYB439* transcript levels were peaking to higher levels than those detected in the WT. This, like the overexpressing lines, had an effect on *KfCCA1* transcript oscillations, with KfCCA1 levels clearly higher at this time (Fig. 4.11). This therefore indicates that a threshold level of *KfMYB439* may be required in order to have this feedback effect on the clock. *KfTOC1* expression levels also fell at 42 h LL, further supporting the role of this *RVE* gene in the core circadian

clock, as all of the clock genes examined were perturbed, although *KfPPCK1* transcript oscillations were still similar to the WT. This could be due to redundancy with other *RVE*s present in *K. fedtschenkoi*.

4.3.3. Gas exchange rhythms

Both *KfMYB439_FL* over-expresser lines had their CO_2 gas exchange rhythms measured using a multi-chamber IRGA system. Both lines had clear perturbations to the circadian rhythm of CO_2 fixation that persists robustly under LL conditions in the WT. *KfMYB439_FL_37C* fixed large amounts of CO_2 before the dark period starts, and stopped fixing CO_2 before the end of the dark and with a more pronounced phase II in LD (Fig. 4.12). *KfMYB439_FL_54B* on the other hand did not start to fix CO_2 until 2 h into the dark, and then carried on until approximately an hour into the light. In addition to these temporal changes in the phase and duration of the bouts of CO_2 fixation, both of these FL lines result achieved much lower total amounts of CO_2 dark CO_2 fixation (Table 4.4). This therefore indicates that there is either a fault with the circadian control of CO_2 fixation, or with stomatal opening, or both. It would therefore be valuable to undertake a detailed study of stomatal conductance in these lines in the future using a IRGA with a clamp-on leaf cuvette such as the LICOR 6400, which is better suited to accurate measurement of stomatal parameters than the CIRAS DC based gas exchange system used here for the LD and LL rhythm analysis. Once these lines had been transferred to LL, both had shorter periods than WT, and their LL rhythms collapsed towards arrhythmia more rapidly with reducing amplitude as the LL experimental run continued (Figs. 4.12 & 4.13).

For line *KfMYB439_RNAi_3C* there was also a marked decrease in nocturnal CO_2 fixation, despite only a small reduction in the target *KfMYB439* transcript level in this

line (Table 4.5). For *KfMYB439_RNAi_10B*, the gas exchange profile was similar to that of the over-expressing lines. CO_2 fixation started at the end of the light, and by the end of the dark fixation has stopped (Fig. 4.15). Phase II was also more pronounced in this line. Both RNAi knockdown lines also displayed a shortened LL period length, and after 7-days in LL, both lines had nearly collapsed into complete arrhythmia (Figs. 4.14 & 4.15). Similar findings were reported previously in Arabidopsis by Mizoguchi *et al*. (2002) who showed arrhythmia of *CCA1/LHY* mRNA levels under LL conditions in *ccal/lhy* mutants. Whilst in *K. fedtschenkoi,* this impact on LL rhythms is not seen at the transcript level in the RNAi lines, it is seen at the level of the CO_2 output rhythm resulting from rhythmic CAM CO_2 fixation.

4.3.4 Metabolite levels

Malate, starch and sugars levels were quantified here due to their central role in the metabolism of the leaves. In particular, malate and starch are the key reciprocating pools of metabolites in a fully functioning CAM leaf, with dark CO_2 fixation requiring PEP generated through starch degradation and leading to malate accumulation, whereas malate decarboxylation in the light leads to recycling of pyruvate to starch. It was therefore important to quantify these metabolites in both the light and the dark in the WT and the transgenic lines. For both the FL and the RNAi lines, there was a trend of them accumulating less malate than the WT in the dark period, and retaining higher levels of malate at the end of the light period (Figs. 4.16 & 4.20). These results are consistent with the gas exchange data, as lower amounts of CO_2 were fixed during the dark in all of these lines, and CO_2 fixation did not commence and curtail coincident with the light – dark and dark- light transitions, but showed either phase advance or delay. This correlated with lower levels of malate

accumulation by dawn, further strengthening the evidence that nocturnal CO_2 fixation associated with CAM was not proceeding optimally in these lines.

All four *KfMYB439* lines retained lower levels of starch than the WT at end of the dark (Figs. 4.17 & 4.21). The highest over-expresser and knockdown lines, namely *KfMYB439_FL_54B* and *RNAi_10B*, both also accumulated lower leaf starch levels at the end of the light. This may be an indication that large perturbations in the transcript levels of *KfMYB439*, either increasing or decreasing, had direct impacts on the efficiency of photosynthesis, and this was certainly evident in terms of the level of dark CO_2 fixation. The other two *KfMYB439* lines had similar levels of starch to the WT at the end of the light. Interestingly, this is also very similar to results for transgenic lines *rNAD-ME* and *rPPDK*, which both showed less malate accumulated at the end of the dark period and more being present at the end of the light. Also, for starch, both the *rNAD-ME* and *rPPDK* lines accumulated less starch by dusk (Dever *et al*., 2015). Furthermore, the more recent, unpublished Illumina RNA-seq data summarised in figure 4.5 revealed clear down-regulation of nearly the whole of the *RVE* family in *rNAD-ME* and *rPPDK*, including *KfMYB439*.

Changes in *KfMYB439* transcript levels also influenced the levels of soluble sugars. These lines had perturbations in sucrose levels, with the over-expressing lines showing a trend towards higher sucrose levels during the light and lower levels in the dark. The RNAi lines show a constant increase in sucrose level. When looking at total accumulation of soluble sugars over a 24 h period, there was seen to be an increase in the average level of sugar during the light, and a decrease during the dark, but this evened out over 24 h to an approximately equal level of sugar in the leaves of the over-expressing lines. The RNAi lines produced contrasting results. *KfMYB439_RNAi_3C*, which only had a small decrease in the transcript level of

KfMYB439, showed similar results to the WT, with only a slight increase in sugar levels (Figs. 4.19 & 4.23). However, *KfMYB439_RNAi_10B* displayed a much greater increase in total sugar amounts (Figs. 4.19 & 4.23). This increase in free soluble sugars may be a mechanism for these plants to maintain their total carbohydrate levels, off-setting the lower levels of starch measured in this lines.

4.3.5 Protein levels and Enzyme activity

The immunoblotting for the *KfMYB439_FL* lines did not reveal any major changes in protein abundances for PEPC, phospho-PEPC and NAD-ME α and β (Fig. 4.24). This was partially supported by the transcript data, where no large differences were detected for the genes encoding these proteins (Figs. 4.6 & 4.7), although *KfPPCK1* transcript levels were higher at both 2 and 6 h into the dark period. However, the phospho-PEPC immunoblot did not indicate that the small elevation of *KfPPCK1* transcript levels was sufficient to achieve a measurable increase in PEPC phosphorylation using this technique.

For the *KfMYB439* RNAi lines differences were detected on the immunoblots for the PEPC dark phosphorylation level and NAD-ME α levels (Fig. 4.25). There was a reduction in phosphorylation at the end of the dark for both RNAi lines. In particular, line *KfMYB439_RNAi_3C* did not show a detectable level of PEPC phosphorylation at the end of the dark (Fig. 4.25B). For NAD-ME α, *KfMYB439_RNAi_3C* showed a large reduction in protein level over the entire 24 cycle, and line *RNAi_10B* displayed reduced levels of NAD-ME α in the light period, but similar levels to the WT in the dark (Fig. 4.25C). These results were not entirely reflected in the transcript abundance data, although *RNAi_10B* did have a lower level of NAD-ME α and NAD-ME β transcripts at 2 h into the light period (Fig. 4.25). The large reduction in the level of

protein detected on the immunoblot for NAD-ME α suggests that a small reduction in *KfMYB439* transcript abundance may be affecting post-transcriptional or post-translational control of NAD-ME α protein levels.

It is important to note though, that protein presence or abundance as determined with immunoblotting does not directly reflect activity. For this reason, the total extractable activity and malate sensitivity of PEPC was determined for the *KfMYB439* transgenic lines,. Both *KfMYB439_FL_37C* and *54B* showed significant increases in activity levels when quantified relative to protein levels in the sample, despite no difference being seen during immunoblotting. When activity was quantified per mg of chlorophyll though, a decrease in activity was seen, which fits better with previous data. This could suggest there is a decrease in the amount of total protein within leaves in KfMYB439, which have resulted in false high activity results. KfMYB439 lines were also found to be much more sensitive to malate inhibition than the WT, both in the light and dark. This suggests that the circadian clock-controlled increase in *PPCK* in the dark, which was at least as large as observed in the WT at both the transcript and PEPC phosphorylation level for the FL over-expresser lines, was no longer able to mediate the WT regulation of PEPC malate sensitivity. The results from the phospho-PEPC immunoblot and the apparent K_i of PEPC for malate measurements do not agree with one another and this is certainly an area that would require further work in order to resolve the details of what has changed in terms of the control of PEPC K_i for malate in the *KfMYB439_FL* over-expresser lines.

The levels of the CAM enzymes measured using immunblots for the *KfMYB439_RNAi* lines did reveal some differences relative to the WT. In particular, decreases in PEPC phosphorylation at the end of the dark could be seen, correlating well with lower apparent K_i of PEPC for malate measured in the desalted extracts

from these lines (Figs. 4.25B & 4.27). NAD-ME α protein levels were also consistently lower for *KfMYB439_RNAi_37C* (Fig. 4.25C). As just discussed for the FL lines, this finding does not correlate with the transcript abundance level for NAD-ME α in this line, suggesting a change in protein regulation or stability may have occurred.

When considering the PEPC activity assays for the *KfMYB439* RNAi lines, both *3C* and *10B* showed an increase in PEPC activity, when expressed per mg protein, but when expressed per mg chlorophyll only *RNAi_10B* showed a significant increase in activity. A decrease in the apparent K_i of PEPC for malate is seen though (Table 4.7). This therefore suggests that either a translational or a post-translational modification may be occurring at the level of PEPC in order to increase its total extractable activity, possibly as part of the counteractive measures these lines take in order to compensate for the reduced apparent K_i of PEPC for malate. However, the protein abundance for PEPC measured with immunoblots for these RNAi lines was very similar to the WT, so these results are again somewhat at odds with the increased total extractable activity.

The *KfMYB439* lines clearly had problems decarboxylating malate, as more malate was present at the end of the light than in the WT; even though less malate was accumulated by the end of the dark period (Figs. 4.16 & 4.20). This could suggest that it may be post-translational modifications that are having the effect on the NAD-ME a protein level, which supports the need to perform NAD-ME enzyme activity assays in the future, as these would reveal whether or not the reduction in NAD-ME α protein level and the increase in end of light malate in the RNAi lines led to and was a consequence of reduced NAD-ME activity in the light in these RNAi lines.

4.3.6 Performance in Drought

A key advantage of the CAM adaptation of photosynthetic CO_2 fixation is that it increases the water use efficiency of the plant. It was therefore important to study the phenotype of the *KfMYB439* transgenic lines under drought-stress.

Both the WT and the transgenic lines all displayed large and significant decreases in fresh weight and moisture content for both above- and below-ground tissues when they were subjected to drought-stress. The WT also showed a significant decrease in dry weight in response to drought for both its shoots and roots (Figs. 4.28 & 4.32). With the transgenic lines, the measured differences for average dry weight in response to drought were not significant for the shoots of the FL lines, but were significant for the roots. Only shoots of RNAi line 3C showed a significant decrease in dry weight with drought, whereas the shoots of 10B and the roots of both RNAi lines all showed a small but none significant decrease in dry weight in response to drought.

Both *KfMYB439_FL* lines showed a significant decrease in fresh and dry weight under well-watered conditions when compared to the WT, showing that over-expression of this RVE family transcription factor impacts on the growth and vegetative yield of the plant. However, there were no significant differences in moisture levels between the FL lines and the WT; suggesting that relative water content was not affected in these lines.

Line *RNAi_3C*, displayed a significant difference in shoot and root fresh and dry weight between well-watered and drought-stress conditions (Fig. 4.33), as was detected in the WT. This does not seem surprising, as this line had a very similar level of *KfMYB439* transcript to the WT. The stronger RNAi knockdown line 10B did not show a significant difference between the two conditions for the dry weight of its shoots or roots, but this line did show a small significant difference in shoot dry

weight under well-watered conditions. By contrast, dry weight of the shoots went up slightly in line 10B under drought stress relative to the WT but these changes were not significant. In the RNAi lines, there was no significant difference for root dry weight between well-watered and drought-stress conditions, although this may be because, on average, the well-watered lines grew less well compared to the WT, resulting in the well-watered and drought-stress values for root dry weight being very similar.

Under prolonged drought-stress, *K. fedtschenkoi* leaves produce and accumulate anthocyanins in their epidermis, which is believed to protect the cells of the leaf from the damaging effects of UV and high-visible light stress (Kolb *et al.,* 2001). Other leaf pigments were also investigated, namely Chlorophyll A, B and Carotenoids, which all play important roles in photosynthetic light harvesting. For the WT and all of the transgenic lines, there was a highly significant increase in anthocyanin levels in response to drought. In particular, line *KfMYB439_FL_37C* had a significantly higher leaf anthocyanin level than the WT in response to drought. The RNAi lines possessed anthocyanin levels close to zero when well-watered, revealing that these lines had lower levels of anthocyanins than the WT when unstressed (Fig. 4.29). However, the RNAi lines were able to produce as much anthocyanin as the WT under the drought-stress conditions (Fig. 4.33).

For the other leaf pigments measured, all showed an increase in levels in respone to drought stress in both WT and the transgenic lines, but there were no significant differences in the level of these photosynthetic pigments when comparing WT levels to those in the transgenic lines under either well-watered or drought-stress conditions. (Figs. 4.30 & 4.34).

When CAM plants are subjected to severe and prolonged drought-stress conditions, a phenomenon referred to as 'CAM-idling' can begin to occur (Luttge, 2004). CAM-idling is defined typified by stomata remaining shut throughout the day and night to conserve water, such that only the CO_2 released from respiration inside the leaf is fixed via PEPC in the dark into malate and, thus, re-cycled (Luttge 2004). This enables malate and starch levels to continue to cycle in a reciprocating manner over the LD cycle, but the daily fluctuations in the malate and starch pools are smaller than those under full CAM conditions. With this 'CAM-idling' physiological response in mind, the malate levels at dawn and dusk in the leaves of the WT and transgenic lines were therefore measured for the well-watered and drought-stressed plants. The malate data for the WT at the end of the dark under both well-watered and drought-stress conditions, indicated that the WT was clearly still conducting CAM, as there was no significant difference in the level of malate accumulated at the end of the dark (Figs. 4.31 & 4.35). In contrast, the drought-stressed FL over-expresser and RNAi knockdown lines showed significant decreases in malate accumulation relative to WT by the end of the dark, suggesting that CAM-cycling was occurring in the *KfMYB439* transgenic lines following this drought-stress treatment. This further strengthens the evidence presented here to support the proposal that the ability to perform optimal CAM was perturbed when *KfMYB439* was mis-regulated. A further possibility was that the changes in *KfMYB439* transcript levels in the transgenic lines enabled the lines to respond to drought in a different way to the WT, as, despite the *KfMYB439* lines having lower fresh and dry weight when well-watered, under drought-stress they did just as well as the WT. This may be because they closed their stomata earlier, suggesting that they could actually have an advantage relative to the WT during very long-term drought. It would be interesting to test this possibility further by performing

a longer drought-stress experiment on these lines, and through making regular measurements of stomatal conductance and transpirational water loss from leaves for both the WT and the transgenic lines throughout the prolonged drought stress treatment.

4.3.7 Reproductive success

The genome of *K. fedtschenkoi* possesses a mutation in the *LEAFY COTYLEDON1* (*LEC1*) gene, which is believed to prevent the developing seed from achieving dehydration tolerance during the final stages of seed maturation (Garces *et al.,* 2007). Thus, *K. fedtschenkoi,* as well as several other plantlet-forming species of *Kalanchoë,* rely on the production of adventitious leaf margin plantlets as their main means of reproduction; a process which is promoted by the *LEC1* mutation in the plantlet-forming *Kalanchoë* species. Thus, the production of plantlets from leaf notches on leaf margins can be used as a measure of reproductive success for *K. fedtschenkoi*. The numbers of available notches on each leaf were counted, as were the number of plantlets formed. The ratio between the number of notches and the number of plantlets provides and indication of how many notches were able to produce a plantlet for each leaf of each line. The *KfMYB439* transgenic lines *FL_54B* and *RNAi_3C* showed significantly lower ratios of notches per plantlet, suggesting they were actually more efficient than the WT at producing plantlets from their leaf margin notches (Figs. 4.36 & 4.37). *KfMYB439_FL_37C* and *KfMYB439_RNAi_10B* were found to have slightly lower ratios still than the WT, but they were not significantly different due to larger levels of variation in the values obtained for these lines. Despite this, these results demonstrate that the *KfMYB439* over-expressers and knockdowns produced a greater number of plantlets than the WT. Thus, although the

KfMYB439 transgenic lines did not perform CAM as efficiently as the WT, they were able to generate more leaf plantlets per leaf notch than the WT, although the ability of those plantlets to establish in soil and produce the next generation was not tested here, so it remains possible that the WT could achieve greater reproductive success than the *KfMYB439* lines because its plantlets establish in soil more efficiently than those of the transgenic lines.

4.3.8 Conclusions

The data presented in this chapter support the conclusion that *KfMYB439* plays a role in the circadian optimisation of CAM. The evidence suggests that this gene functions closely to, but does not play a role in the core of the clock. *KfMYB439* mis-regulation had a widespread, but relatively small, effect on the transcript abundance of several key CAM-associated genes. Gas exchange analysis revealed that the transgenic lines had phase changes in the timing of their CO_2 fixation in the dark, and also that they did not achieve as much CO_2 fixation as the WT over the entire dark period. Changes to this gene also impacted on the amount of starch and malate present in leaves at the end of the dark period, with both metabolites being reduced. This was consistent with the finding that the apparent K_i of PEPC for malate was reduced relative to the WT in the transgenic lines, and the level of phospho-PEPC in the dark was also reduced relative to WT in the RNAi lines, which correlated well with the finding that these lines did not fix as much CO_2 as the WT in the dark. Furthermore, under optimal, well-watered conditions, the WT achieved a higher dry weight yield than the transgenic lines, especially the FL over-expresser lines, supporting the proposal that the WT was able to grow more efficiently that the transgenic lines. However, following drought-stress, both the WT and the transgenic lines performed equally

well in terms of dry weight yields for roots and shoots. Interestingly, the RNAi lines performed slightly better under drought than the WT. Furthermore, the *KfMYB439* transgenic lines were able to produce adventitious leaf margin plantlets in a higher proportion of their available leaf notches suggesting they may be able to achieve a higher reproductive output and thus fitness, although the establishment of the plantlets will need to be studied to confirm this.

The goal of this work was to understand the role and function, if any, of *KfMYB439* in the regulation of CAM in *K. fedtschenkoi*. An important consideration when developing understanding of the function of this gene relates to the fact that it is a member of the *RVE* family of single MYB-repeat transcription factors. Functional redundancies and interactions have been demonstrated for several of the *RVE* genes in Arabidopsis (Zhang *et al.,* 2007; Farinas *et al.,* 2011; Hsu *et al.,* 2013). It is therefore clear from this paradigm established in Arabidopsis that knocking down one *RVE* may be compensated for by other members of the family. It will be interesting in the future to investigate the regulation of the other *RVE* genes in the *K. fedtschenkoi* genome in these *KfMYB439* transgenic lines using RT-qPCR to quantify the transcript abundance of the other RVEs over the LD cycle and under LL free-running conditions. Also, as single MYB-repeat TFs like *KfMYB439* are known to form and function as heterodimers in Arabidopsis, the phenotypic effects in the transgenic lines studied here may have been lessened because if the other *RVE* family TFs were not present at the correct level for efficient functional dimar formation, then the downstream functions and impacts of *KfMYB439* and its functional partner(s) would not be evident. It will therefore be important for future work to identify the *RVE*(s) that interact directly with KfMYB439 and generate transgenic *K. fedtschenkoi* lines in which both RVEs are either over-expressed or silenced with RNAi.

Another key area for future work on *KfMYB439* will be to identify its target sites within the genome in order to understand how this TF binds to and regulates the expression of the downstream genes. A valuable technique for gaining an insight into the genes targeted by KfMYB439 would be Chromatin Immunoprecipitation followed by sequencing (ChIP-seq), which allows the identification of all of the promoters bound by a particular TF by capturing the TF-DNA complexes and then sequencing the DNA regions of the genome that were bound by the TF. To this end, Haemaglutenin (HA)-tagged *KfMYB439* transgenic lines of *K. laxiflora* were generated as part of this project, but unfortunately time ran out before it was possible to attempt the ChIP-seq experiments to identify the genomic targets of this TF. This will be an exciting area for future work, as it will reveal on a genome-wide scale, which genes, and thus biological processes, are being directly regulated through the binding of *KfMYB439* to their promoters.

Chapter 5 – Characterisation of the CAM-associated phenotypic changes in transgenic RNAi and over-expression lines of *K. fedtschenkoi* with altered expression levels of the CAM-induced and clock-controlled transcription factor *CAM-INDUCED bZIP1* (*KfCIB1*)

5.1 Introduction

CAM-INDUCED bZIP (*KfCIB1:* KF13220 in the *K. fedtschenkoi* draft genome) was identified previously through semi-quantitative RT-PCR analysis, plus quantitative RNA-seq analysis, as being strongly CAM-induced and regulated by the circadian clock, as described previously in Chapter 3 (Hartwell *et al.,* Unpublished data). *KfCIB1* was characterised as displaying a pre-dawn peak in its transcript abundance, similar to the CAM-induced and clock-controlled protein kinase *KfPPCK1*. Due to this transcript abundance profile, *KfCIB1* was selected for detailed analysis *in planta* as a candidate transcription factor that could play a role in linking the circadian clock output signal transduction cascade to CAM.

5.1.1 bZIP transcription factors

The bZIP superfamily of transcription factors has been shown to interact in a network of homo- and heterodimers to regulate a diverse set of cellular processes in all eukaryotes. bZIP regulated processes include cell survival, lipid metabolism and even cancer progression. Plant bZIP TFs, which are one of the largest and most diverse families of plant TFs, make up 4 % of all known plant TFs and regulate a wide range of crucial processes, including abiotic stress responses (Bartels *et al.,* 2005), nitrogen and carbon metabolism (Vincentz *et al.,* 2003), energy metabolism (Baena-Gonzalez *et al.,* 2007), and tissue differentiation (Gibalova *et al.,* 2009). These TFs are represented by large families in the plant genomes that have been published to date. For example, there are 77 bZIPs in Arabidopsis, 89 in black cottonwood, 92 in rice, 55 in grapevine and 131 in soybean (Jakoby *et al.,* 2002; Liu *et al.,* 2014)

5.1.1.2 bZIP structure

All bZIP TFs consist of two subdomains; a basic region, which contains a nuclear localisation signal and a sequence-specific DNA binding site, and an alpha-helical leucine zipper, which contains varying numbers of heptad repeats, enabling dimerisation to occur (Fig. 5.1). The Leucine (Leu) zipper often contains Leu amino acids, but can also contain other bulky hydrophobic residues such as Valine, Phenylalanine (Phe) or Isoleucine (Ile) leading to dimerisation specificity (Liu *et al.,* 2014).

Figure 5.1 Primary bZIP protein structure. bZIPs consist of two domains; basic region (in blue) and the zipper region (in grey). Highly conserved regions are shown in the dark colour boxes. Image taken from (Jakoby *et al.,* 2002).

Other conserved domains in addition to the bZIP domain have been found, which function as transcriptional activators. They have been shown to be highly conserved between groups and subfamilies, as have the intron/exon structures (Liu *et al.,* 2014). This suggests purifying selection has occurred on this family of transcription factors during evolution.

For plant bZIPs to bind to DNA, they must form a dimer. bZIPs are able to form homodimers, heterodimers and also quasi-homodimers, and these interactions are controlled by a variety of post-translational modifications, such as phosphorylation, protein charge, interaction with other proteins, and available partners present in a tissue at various times of day (Hurst 1995; Lee *et al.,* 2006; Schutze *et al.,* 2008).

Plant bZIP proteins have been shown to bind preferentially to DNA sequences with upstream promoter ACGT elements at their core (Foster *et al.,* 1994). Binding specificity is then defined by the flanking sequence of the DNA, but plant bZIPs have been shown to commonly bind to the A-box (TACGTA), C-box (GACGTC) and G-box (CACGTG), although not all binding sites have been shown to be palindromic (Izawa *et al.,* 1993; Choi *et al.,* 2000; Fukazawa *et al.,* 2000).

5.1.2 Functional diversity of bZIPs

As already stated, plant bZIPs regulate a wide range of important processes (Table 5.1). Grouping of these bZIPs is dependent on the similarity between basic regions, conserved motifs and also the size of the leucine zipper (Jakoby *et al.,* 2002). More emphasis was put onto conserved motifs to help with role determination when bZIP groups were defined.

Table 5.1 Largest bZIP groups and their commonly known functions.

Group	Functions
A	Abscisic acid (ABA) and stress signalling
C	Seed Storage, Pathogen responses, Response to environmental change
D	Pathogen defence, plant development
G	Light signal transduction
H	Photomorphogenesis
I	Vascular development
S	Carbohydrate metabolism, stress responses, energy use

5.1.3 *KfCIB1*

Many plant *bZIP* genes have been well classified in model species, such as Arabidopsis and crops including rice and barley, but in other species, very little is known about the family, including in Kalanchoe.

This TF was originally identified as a CAM-induced and circadian clock controlled TF in the salt and drought-stress inducible facultative CAM species *M. crystallinum* (Boxall *et al.,* Unpublished data). An orthologous bZIP was subsequently identified amongst 5000 Sanger sequenced expressed sequence tags recovered from a traditional cDNA library generated using a dark sample of CAM leaf RNA from the obligate CAM species *K. fedtschenkoi*. Earlier work by a previous PhD student in the Hartwell Lab in Liverpool had investigated the light/ dark and circadian regulation of this *KfCIB1* gene during leaf development, and had studied possible promoter motif DNA-binding sites that this TF may interact with in candidate downstream target genes associated with CAM (Meszter, 2010). This work also characterised the *KfCIB1* Open Reading Frame (ORF) and found that it encoded a basic domain and also possesses 8 heptad repeats, which is characteristic of group S bZIPs (Meszter 2010). This was confirmed using phylogenetic analysis of multiple sequence alignments and the construction of phylogenetic trees. It was shown that all members of group S possessed, like *KfCIB1*, 8 heptad Leu repeats.

More recently, *de novo* sequencing and assembly of the *K. fedtschenkoi* genome and transcriptome using 454 and Illumina sequencing, and accompanying quantitative SOLiD RNA-Seq data cataloguing the light/ dark regulation of gene transcript levels in both C_3 and CAM leaves of *K. fedtschenkoi,* confirmed that *KfCIB1* was amongst the most strongly induced TFs in CAM leaves compared to C_3 leaves (*KfCIB1* was

induced over 60-fold in CAM leaves relative to C_3 leaves at its temporal peak in the middle of the dark period; Boxall, Dever, Knerova and Hartwell, unpublished data).

5.1.3.1 *KfCIB1* in *K. fedtschenkoi*

As mentioned above, *KfCIB1* levels increased ~60-fold in CAM performing leaves (LP6) compared to C_3 leaves (LP1) (Hartwell *et al*., Personal communication). Previous semi-quantitative RT-PCR and RNA-seq datasets had demonstrated that this TF was under light/ dark control at the level of its transcript abundance, peaking in the middle of a 12 h dark period, and was also under circadian clock control in LL free-running conditions (Chapter 3).

Comparisons were then conducted between *KfCIB1* and group S bZIPs with known functions. The highest similarities and functions were shown to be in abiotic stress responses; including salt, drought and dehydration (Meszter 2010).

Furthermore, it had been found using electrophoretic mobility gel-shift analysis that heterologously-expressed *KfCIB1* had the potential to bind to the *KfPPCK1* promoter at three potential ACGT motifs (Meszter, 2010). Furthermore, it was speculated in this PhD thesis that, as the timing of *KfCIB1* transcription matched the phase of *KfCCA1* oscillations, then *KfCCA1* could be directly modulating *KfCIB1* (Meszter, 2010). A similar link has been reported already between *AtCCA1* and a group S bZIP (*AtbZIP1*) in Arabidopsis and also AtCCA1 and bZIP ELONGATED HYPOCOTYL5 (HY5), which have been shown to physically interact. (Andronis *et al.,* 2008; Krouk *et al.,* 2008).

All of this data therefore led to *KfCIB1* being identified as a transcription factor that was a strong candidate to be involved in the circadian optimisation of CAM.

Transgenic RNAi knockdown and constitutive over-expression lines of *K.*

fedtschenkoi were therefore generated in order to understand more about the *in planta* functions of this TF. After initial screening (Chapter 3), the two best over-expresser and two best RNAi transgenic lines were chosen to be worked on in detail, as discussed in Chapter 3.

5.2 Results

5.2.1 Bioinformatic Analysis of *KfCIB1*

5.2.1.1 Identification of conserved domains

The highest BLAST hit from the *Arabidopsis* TAIR website was for another bZIP transcription factor: BASIC LEUCINE-ZIPPER 42 (ATbZIP42; Arabidopsis accession number: At3g30530). The bit score for this result was 56, with an E value of 7e-07.

After translating and aligning KfCIB1 and ATbZIP42, both show conserved sequences around the basic domain and heptad-repeat Leucine zipper regions (Fig. 5.2A). All of the predicted conserved amino acids in the basic region for the DNA binding site were conserved between these two amino acid sequences, but within the leucine zipper, the two sequences diverged due to a number of amino acid substitutions. In the heptad repeats of the leucine zipper domain, every seventh amino acid is predicted to be a Leucine. AtbZIP42 possessed a Leucine residue at every site in the heptad repeat. However, for KfCIB1 there were several amino acid changes in this region. The blue stars show Leucines found just in AtbZIP42, and the red show amino acids found in both sequences (Fig. 5.2B). These amino acid changes are likely to affect dimerization, and so *KfCIB1* may have taken on new roles, or new interactions compared to *AtbZIP42*.

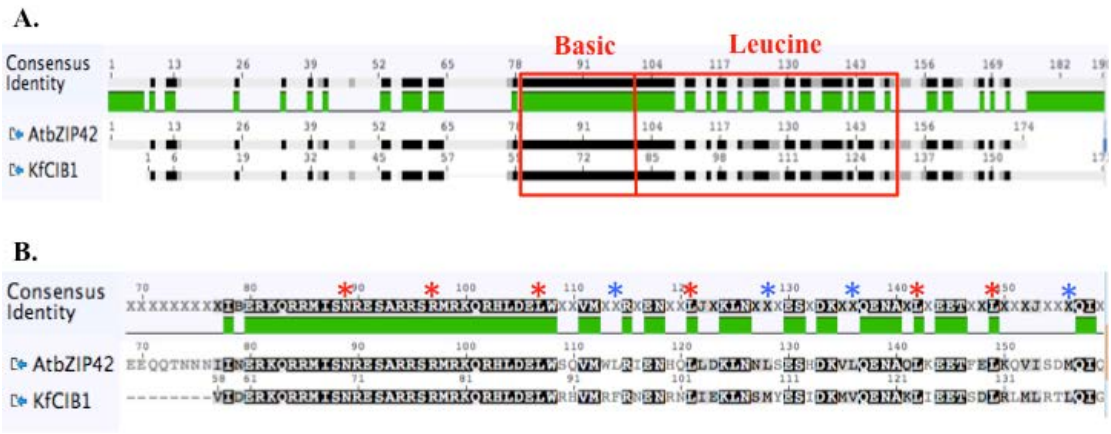

Figure 5.2. Geneious global amino acid alignment of KfCIB1 with its closest ortholog in Arabidopsis: AtbZIP42. **A.**There is 100 % conservation in the basic domain involved in nuclear localisation and DNA binding, but less conservation in the Leucine region, where the heptad repeats of Leucines have not been as well conserved in KfCIB1. **B.** Conserved regions found in many bZIPs and also between the two sequences have a red star. Blue stars represent heptad repeat Leucines, which were not conserved in KfCIB1, but were highly conserved in bZIP42 (and in other bZIP TFs).

5.2.1.2 Perturbation of the regulation of *KfCIB1* in transgenic lines of *K. fedtschenkoi* in which key CAM-associated and clock-associated genes were mis-regulated

Transgenic lines of *K. fedtschenkoi* have been generated previously in the Hartwell lab in which a range of CAM- and clock-associated genes have been either down-regulated with RNAi, or expressed constitutively. Four of these transgenic lines (*rNAD_ME1*, *rPPDK*, *rPPCK1_3* and *TOC1_OX*) have recently had their transcriptome-wide responses to the mis-regulation of the targeted gene assayed using Illumina RNA-seq. These transgenic lines all displayed large perturbations in their daily pattern of CAM CO_2 fixation and the circadian rhythms of CO_2 fixation in LL conditions. The RNA-seq dataset therefore allowed the regulation of *KfCIB1* to be studied in transgenic lines that have large changes in CAM physiology.

The first noteworthy finding from this RNA-seq dataset was that the regulation of *KfCIB1* transcript levels in the wild type did not match the initial SOLiD RNA-seq dataset for C_3 and CAM leaves in terms of the temporal timing of the peak of *KfCIB1* transcript levels. In the original SOLiD RNA-seq dataset, *KfCIB1* trasncript levels peaked at 18:00 (middle of the 12 h dark period), whereas the more recent Illumina RNA-seq dataset revealed a temporal peak for *KfCIB1* at 22:00 (2 h before dawn). *KfCIB1* also peaked at 22:00 in semi-quantitative RT-PCR analysis presented in this chapter. Based on these findings, the wild type daily peak of *KfCIB1* transcript abundance is hereafter assumed to occur at 22:00 (2 h before dawn).

When comparing *KfCIB1* regulation between the wild type and the four transgenic lines, the down-regulation of *KfPPCK1* in line *rPPCK1_3* did not change *KfCIB1* transcript levels (Fig. 5.3). *TOC1_OX* led to a > 50 % decrease in the transcript level of *KfCIB1* at its pre-dawn peak, which concurs with the fact that *KfCIB1* was regulated by the circadian clock, as over-expression of a core clock genes such as TOC1 is known to cause arrhythmia of clock-regulated outputs. In line *rNAD-ME1,* the transcript abundance level of *KfCIB1* increased throughout the dark period. The daily peak of *KfCIB1* occurred 4 h earlier than the wild type in *rPPDK1*. *KfCIB1* transcript levels were also elevated in line *rPPDK1,* peaking 4-fold higher than the wild type at 22:00 (Fig.5.3). This up-regulation of *KfCIB1* in the *rNAD_ME1* and *rPPDK* lines of *K. fedtschenkoi,* which both have a large, almost complete, reductions in their ability to perform dark CO_2 fixation associated with CAM, supports the original hypothesis that this gene plays an important role in the temporal coordination of the daily CAM cycle. Furthermore, the fact that the two transgenic lines lacking key enzymes in the decarboxylation pathway, which releases CO_2 from malate in the

light period, had elevated levels of *KfCIB1* suggests that *KfCIB1* may play a specific role in the regulation of the CAM decarboxylation pathway.

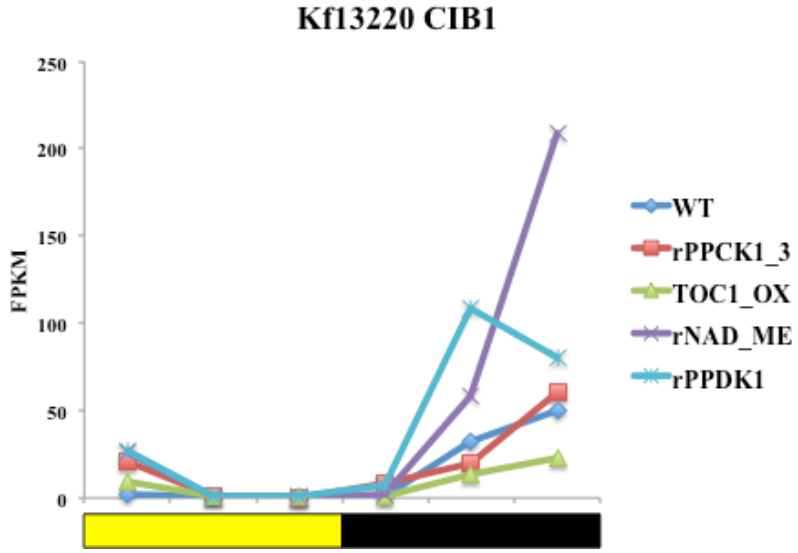

Figure 5.3. Illumina RNA-seq transcript abundance (FPKM) data showing the expression of *KfCIB1* transcript in wild type (WT) and transgenic lines: *KfPPCK1* RNAi knockdown (*rPPCK1_3*), TOC1 over expresser (*TOC1_OX*), *NAD-ME β* knockdown (*rNAD_ME*) and *PPDK* knockdown (*rPPDK1*). Samples were collected in 12:12 LD at 2, 6 and 10 hours into both light and dark *KfCIB1* transcript levels were reduced in *TOC1_OX*, whilst both *rPPDK* and *rNAD-ME* showed large increases in *KfCIB1* levels during the dark period. *KfCIB1* levels were very similar to the wild type in line *rPPCK1_3*.

5.2.2 The impact of the transgenic perturbation of *KfCIB1* levels on the light/dark regulation of key CAM- and circadian clock-associated genes

Plants were entrained to 12:12 L:D conditions (Materials and Methods 2.1.2), and LP6 was sampled in biological triplicate every 4 hours over a 24 h period. A large range of genes, as well as *KfCIB1,* were screened using semi-quantitative RT-PCR, including the core clock genes *KfCCA1* and *KfTOC1*, CAM carboxylation pathway genes *KfPPCK* and *KfPEPC*, CAM decarboxylation pathway genes *KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD- ME β*, and also genes associated with starch metabolism and the sub-cellular transport of sugars *KfGPT2, KfMEX1* and *KfGWD.*

5.2.2.1 Light/ dark regulation of CAM and clock genes in *KfCIB1* full-length overexpresser lines

For both *KfCIB1_FL_5B* and *11B,* there were large increases in the transcript abundance level of *KfCIB1* relative to the wild type, with a 4.3-fold and 1.4-fold average increase in transcript abundance in the two lines respectively (Fig 5.4; Fig 5.5). Although none of the genes assayed showed transcript abundance changes of the magnitude of the change in *KfCIB1* transcript levels, in both lines, there were still noticeable differences in the daily oscillations of the transcript levels of the central circadian clock genes *KfCCA1* and *KfTOC1*. In both lines *KfCCA1* shows decreases in expression around dusk and the dark period, whilst *KfTOC1* only shows a decrease in *FL_5B* at peak expression (Figs. 5.4 & 5.5). *FL_11B* on the other hand shows an increase at peak expression time. Several key CAM-associated genes also show expression changes (Figs. 5.4 and 5.5). *KfCIB1_FL_5B* and *11B* both showed transcript abundance differences for the genes encoding the two subunit of NAD-ME, namely *NAD-ME α* and *NAD-ME β*. *KfCIB1_FL_5B* showed stronger light/ dark oscillation in the transcript abundance of these two genes compared to the wild type (Fig. 5.4), whilst *KfCIB1_FL_11B* showed elevated levels of *NAD-ME α* and *NAD-ME β* throughout the 24 h cycle, with a small peak in the middle of the light period (Fig. 5.5). Furthermore, levels of *KfPPDK* and *KfPPDK-RP* displayed small perturbations in their abundance in both transgenic lines, with more transcript present during the light period, and less is in the dark. For *KfCIB1_FL_5B,* and more so *FL_11B* , *KfGWD1, KfMEX1* and *KfGPT2* also showed a higher amount of transcript during the day. Both lines also show an increase in *KfPEPC* expression, with 5B showing increases in the light, and 11B having increased expression throughout most of the day.

Other genes show differences in transcript level, but in only one line. This seems reasonable though, due to the fold differences in the level of *KfCIB1* over-expression, which could indicate a threshold level is required for an effect. Notably, *KfPPCK1* transcript levels were reduced relative to the wild type during the first 6 hours of the dark period in *KfCIB_FL_5B*, which had the larger fold change of *KfCIB1* expression (Fig. 5.4).

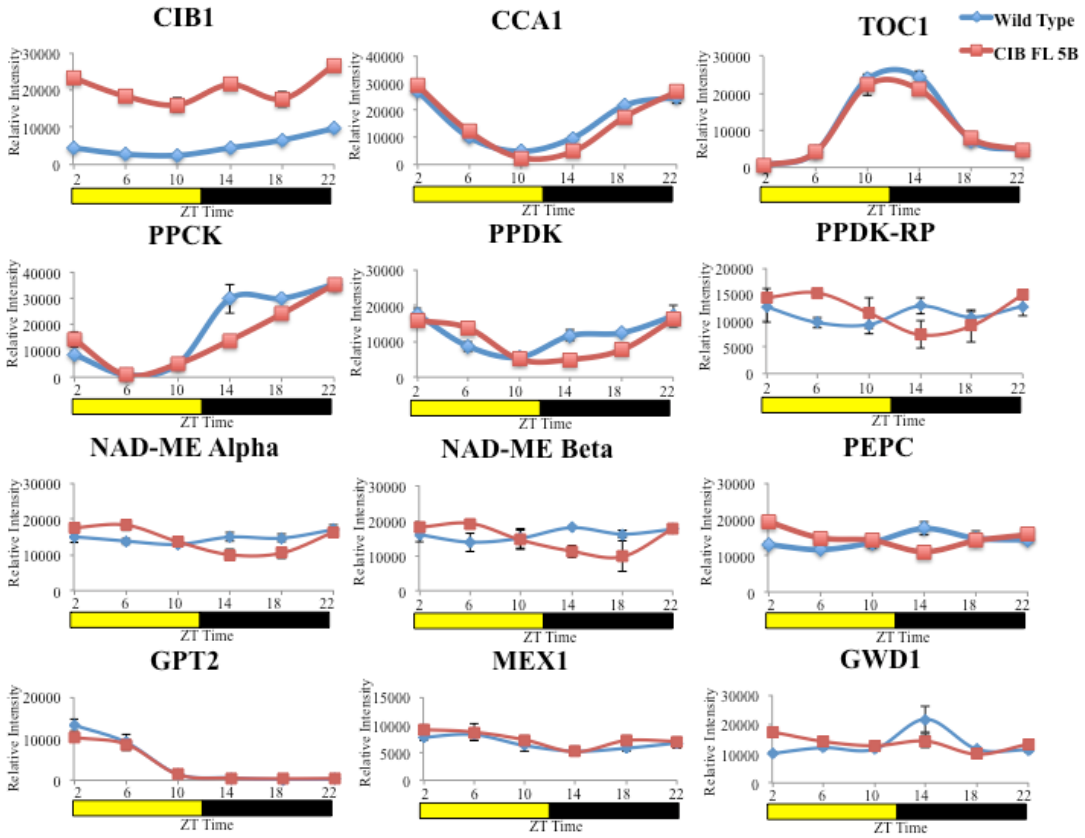

Figure 5.4. Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCIB1_FL_5B* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line *FL_5B*. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

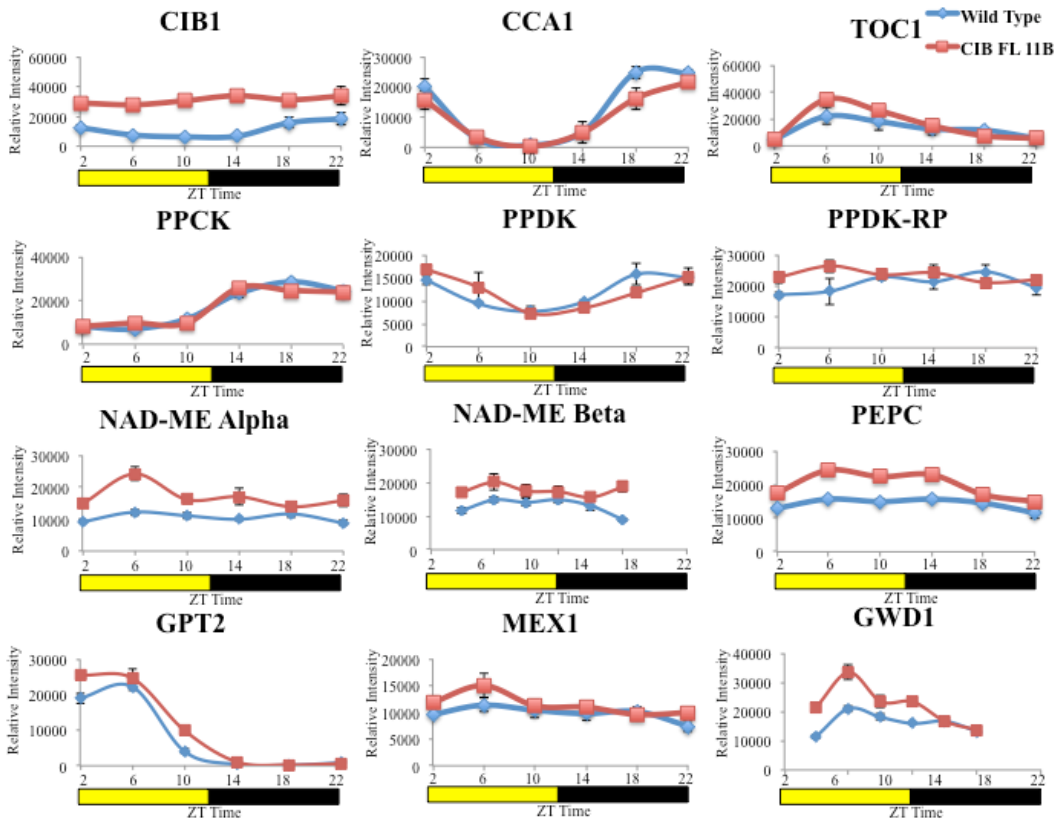

Figure 5.5 Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCIB1_FL_11B* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line *FL_11B*. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

5.2.2.2 Light/ dark regulation of CAM and clock genes in *KfCIB1* RNAi knockdown lines

KfCIB1_RNAi_12A and *30A* were also analysed using semi-quantitative RT-PCR (Figs 5.6 and 5.7). *KfCIB1_RNAi_12A* showed a 118.7 fold decrease at 10h dark (22:00) compared to wild type and 30A showed a 9.6 fold decrease (Fig 5.6; Fig 5.7). As with the *KfCIB1_FL* over-expressers, there were only small differences in the transcript levels of the clock genes for either line, but the RNAi lines appear to show small increases in clock gene expression, whereas in the overexpressing lines, a small decrease was seen. Interestingly, reducing the transcript abundance of *KfCIB1* also had only a small effect on the transcript abundance of the CAM genes screened. Small differences in transcript levels were detected for *NAD-ME α, NAD-ME β, PPDK-RP* and *PEPC,* but compared to the differences in the FL over-expressor lines, these were generally smaller changes. One interesting and consistent findings concerned the transcript level of *KfPPCK1*, which was elevated slightly at its pre-dawn peak in both RNAi lines, and showed a decrease in both *KfCIB1_FL*s.

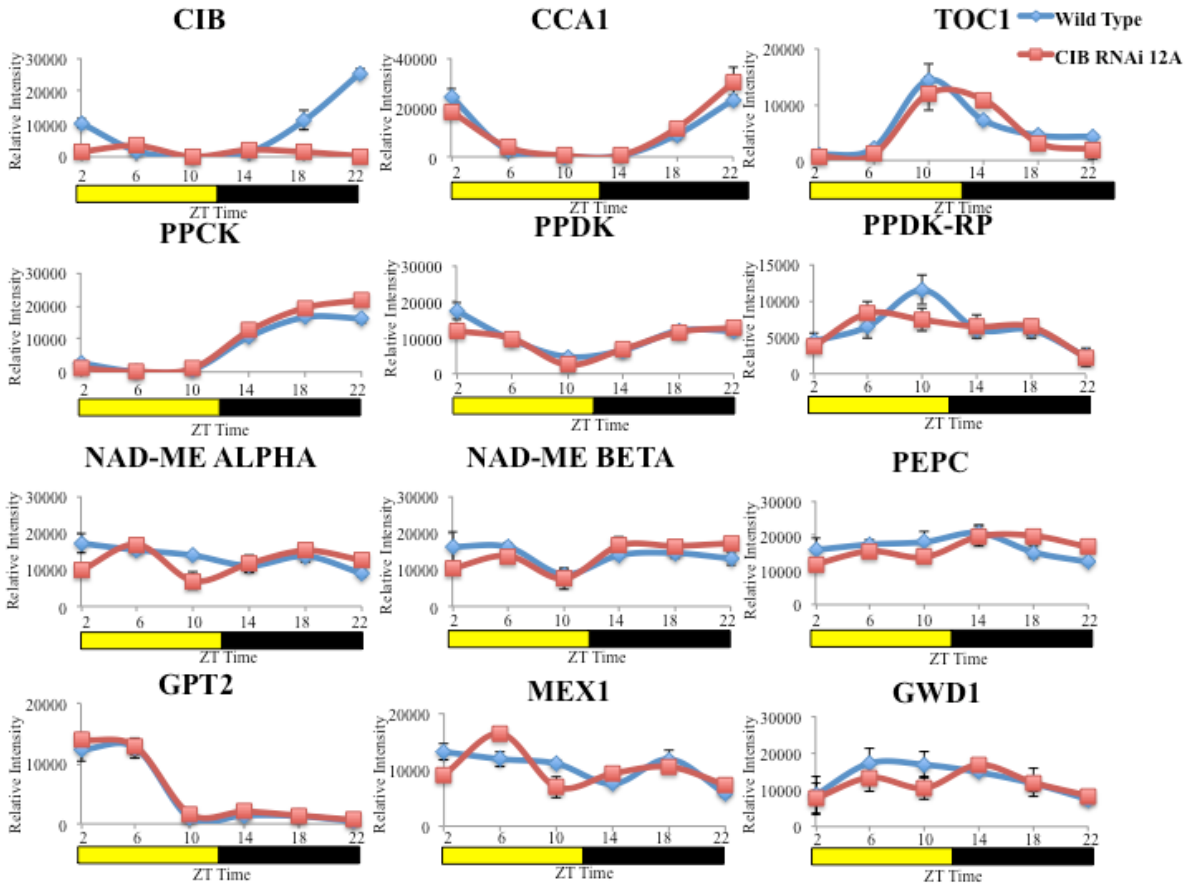

Figure 5.6 Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCIB1_RNAi_12A* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line *RNAi_12A*. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

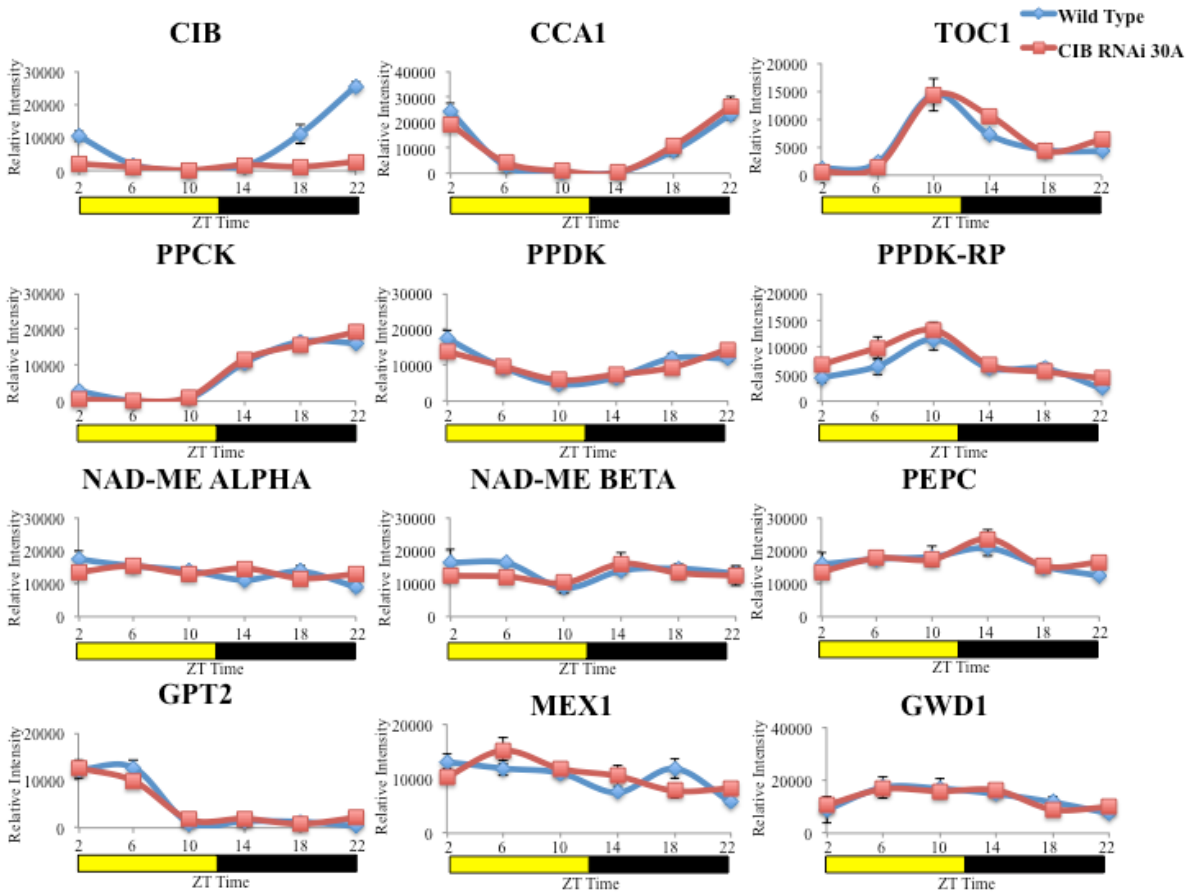

Figure 5.7 Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCIB1_RNAi_30A* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line *RNAi_30A*. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

5.2.3. The impact of perturbing *KfCIB1* levels on circadian oscillations in the transcript abundance of circadian clock controlled genes

KfCIB1 itself had already been confirmed as being under circadian clock control prior to the original generation of the transgenic lines which pre-dated the start of this project. *KfCIB1* was shown to continue to oscillate at reduced transcript levels under in constant light (LL) conditions (Fig. 3.1B). It was therefore important to investigate whether perturbations in *KfCIB1* transcript levels in the transgenic lines had any effect on the plant's ability to maintain oscillation within the central circadian clock and the output pathway to CAM. For LL time courses, only one over-expresser and RNAi knockdown was sampled. A 10 h Light and 10 h Dark sample of LP6 was collected in biological triplicate at the start of the time course where the plants were still under the 12:12 light/ dark cycles used for pre-entrainment. The plants were then transferred to constant light, temperature and humidity conditions (LL). They were then left in these conditions for over 24 h to enable them to enter true free-running conditions prior to sampling. Samples were then collected every 4 h over 3 days, starting at 26 h after the change to LL.

5.2.3.1 The impact of perturbing *KfCIB1* levels on circadian oscillations in *KfCIB1* full-length overexpression lines

KfCIB1_FL_11B was chosen for circadian analysis. The level of overexpression was consistently higher than the wild type in both LD and LL conditions, supporting the prior screening (Chapter 3), and LD time course data (Figs. 3.1B and 5.5). These results demonstrate that, even though the CaMV35S promoter was used to drive the over-expression of the *KfCIB1* full-length open reading frame, circadian control was still exerted on the transcript abundance (Fig. 5.8). Peak *KfCIB1* expression in LL can be seen to be phase advanced by approximately 4 h relative to the wild type, and the *KfCIB1_FL* line was able to continue to peak at the end of subjective dark, despite being in constant free-running

conditions. In the WT, the transcript abundance of *KfCIB1* peaked at the start of the subjective light period, 4 h after the peak in the *KfCIB1_FL* line (Fig. 5.8).

The clock gene *KfCCA1* continued to oscillate in both WT and *KfCIB1_FL*, and until 38h into LL, the two lines oscillate with similar abundances and phases (Fig. 5.8, CCA1 – first white dotted box). After this the transgenic lines increases expression, and for many time points after 38h LL, *KfCCA1* shows higher expression in *KfCIB_FL_11B*.

Interestingly for *KfTOC1,* both lines no longer showed peaks at their expected times after 24 h under LL, and *KfCIB_FL_11B* was four hours out of phase with wild type for its *KfTOC1* transcript oscillations (Fig. 5.8). *KfPPCK1* oscillations were very similar in the WT and *KfCIB1_FL*, except that the amplitude of the transcript oscillations dampened more rapidly in the *FL* line towards the end of the 3 day time course.

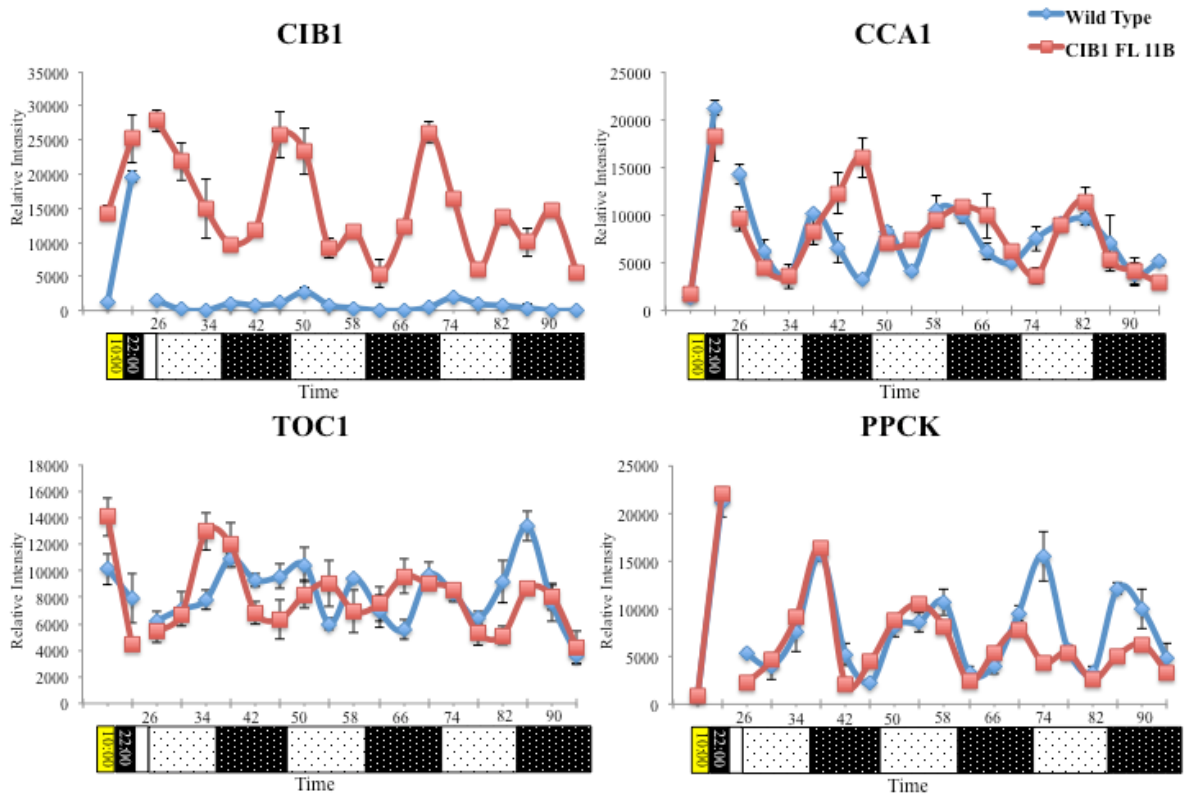

Figure 5.8 LL Constant light time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (Blue) and transgenic line *KfCIB_FL_11B* (Red) measured in full CAM leaves (LP6) sampled at 4h intervals over a LL light regime. The first two samples collected were in LD and then the light regime was turned to LL. Plants were left for 26h in LL before sampling commenced. Fluctuations in the transcript abundance of clock genes (*KfCCA1* and *KfTOC1*) and CAM gene *KfPPCK* were determined, along with *KfCIB1* expression, for both Wild type and *KfCIB_FL_11B*. The yellow bar represents light period, the black bar represents the dark period, and the white bar represents a break in sampling for the first 26h in LL. White boxes containing black dots represent subjective light and black boxes with white dots represent subjective dark. To entrain *KfCIB_FL_11B*, The plants were entrained to lights on at ZT0 (00:00 h), and went off at 12:00h. Light sample was collected at 10:00 and dark sample was collected at 22:00. LL samples were collected from 26:00 in LL every 4 h for three days.

5.2.3.2 The impact of perturbing *KfCIB1* levels on circadian oscillations in *KfCIB1* RNAi knockdown lines

When *KfCIB_RNAi_30A* was assayed in LL constant conditions, the effect of the RNAi interference seemed to be overridden. In the initial light and dark samples collected it can be seen that *KfCIB1* transcript level was reduced, as it was in the previous LD time course, but after 26 hours in LL, the level of *KfCIB1* transcripts was higher than the WT (Fig. 5.9). The *KfCIB1_RNAi_30A* line appears from this experiment to over-express the *KfCIB1* transcript at many of the remaining time points throughout the remainder of the 3-day LL time course. Interestingly, again *KfCIB1* displayed a 4 h phase advance in its transcript peak relative to the wild type and still peaked at the end of the 'subjective dark', until 82 h LL (Fig. 5.9 – third white dotted box). The wild type began to peak at the start of the subjective light very soon after the start of the LL time course. After 42 h, *KfCCA*1 began to show differences in transcript abundance, such that the *KfCCA1* transcript abundance results for *KfCIB1_RNAi_30A* were very similar to the results for the *KfCIB1_FL_11B* over-expresser under LL (compare Figs. 5.8 & Fig. 5.9).

KfTOC1 transcript levels also displayed differences between the wild type and *KfCIB1_RNAi_30A* after 42 h under LL (Fig. 5.9 – first black dotted box). For *KfCCA1*, it appears that whilst wild type possesses a shorter period, *KfCIB1_RNAi_30A* becomes arrhythmic, as there is very little cycling seen past the 42 h point. *KfPPCK* levels again do not differ much from wild type, as reported above for *KfCIB1_FL_11B*, except that by the last subjective light and dark period there was a peak at the end of the subjective light in the RNAi line, similar to the results for the *KfCIB1* over-expresser line (Fig. 5.9).

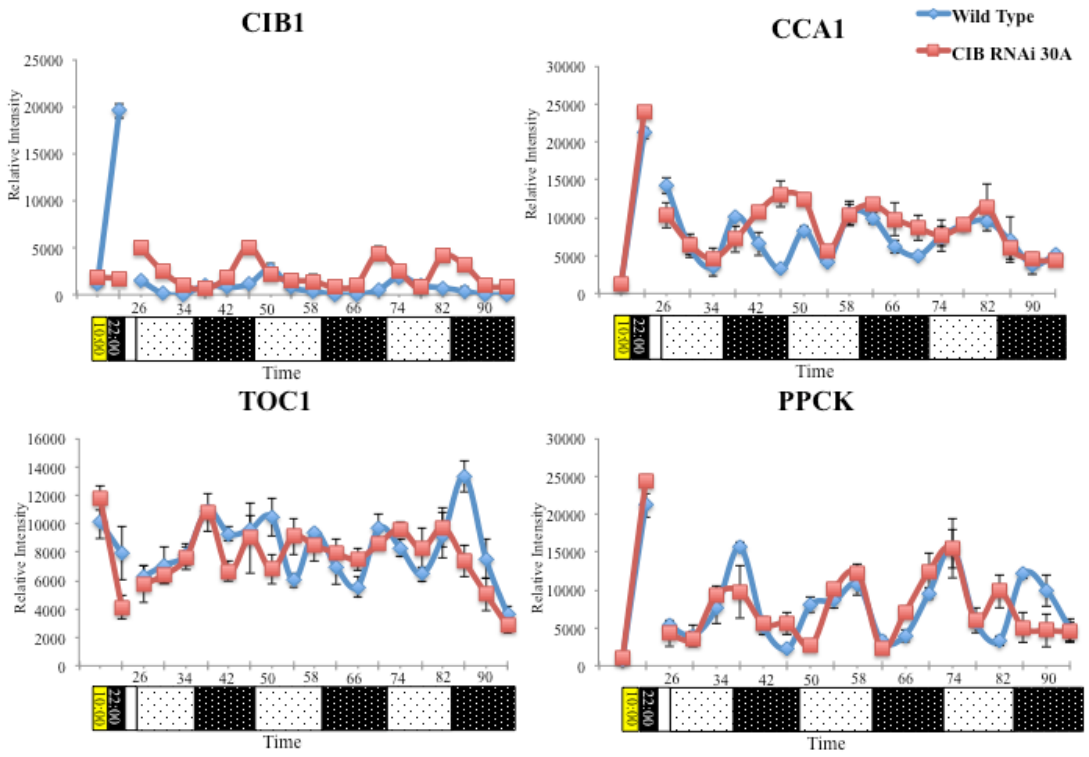

Figure 5.9 Constant light time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (Blue) and transgenic line *KfCIB_RNAi_30A* (Red) measured in full CAM leaves (LP6) sampled at 4h intervals over a LL light regime. The first two samples collected were in LD and then the light regime was turned to LL. Plants were left for 26h in LL before sampling commenced. Fluctuations in the transcript abundance of clock genes (*KfCCA1* and *KfTOC1*) and CAM gene *KfPPCK* were determined, along with *KfCIB1* expression, for both Wild type and *KfCIB_RNAi_30A*. The yellow bar represents light period, the black bar represents the dark period, and the white bar represents a break in sampling for the first 26h in LL. White boxes containing black dots represent subjective light and black boxes with white dots represent subjective dark. To entrain *KfCIB_RNAi_30A*, The plants were entrained to lights on at ZT0 (00:00 h), and went off at 12:00h. Light sample was collected at 10:00 and dark sample was collected at 22:00. LL samples were collected from 26:00 in LL every 4 h for three days.

5.2.4 The impact of alterations in *KfCIB1* transcript levels in transgenic *K. fedtschenkoi* on CO_2 exchange rhythms under LD and LL cycles

5.2.4.1 Impact of overexpressing *KfCIB1* on diurnal gas exchange rhythms

5.2.4.1.1 Gas exchange measurements on overexpressing lines using the multi-channel IRGA

KfCIB1_FL_5B and *11B* were both analysed using the multi-chamber IRGA system in order to study their CAM CO_2 fixation rhythms. Leaf gas exchange measurements allow the physiology of the transgenic lines to be studied in detail in order to determine whether or not the changes in the target gene influence the control of CO_2 fixation associated with CAM. Under LD conditions, the results were very interesting, with large spikes of CO_2 fixation and respiratory CO_2 release evident for the transgenic lines (Figs. 5.10 & 5.11). From when lights turn off, and for the first 1-2 h of darkness, large spikes of CO_2 were released from the CAM leaves of line *KfCIB1_FL_5B* (Fig. 5.10). In addition, there were also large spikes of CO_2 fixation at the start of the light period, which was phased such that it could be interpreted as a highly exaggerated phase II, compared to the wild type (Fig. 5.10A & B). With *KfCIB1_FL_11B*, only one CO_2 spike can be seen, which is not as consistent with *KfCIB1_FL_5B*, and instead shows CO_2 loss at the point of light conditions changing to either light or dark (Fig. 5.11A & B). Throughout the light period in LD both *KfCIB1_FLs* appear to be constantly switching between fixing and releasing varying amounts of CO_2, as the CO_2 fixation trace for these lines was constantly fluctuating, although to different extents. By contrast, the trace for the wild type leaves was comparatively stable and smooth even though a single infra-red gas analyser makes all of the CO_2 differential measurement in this gas exchange system, which suggests that these rapid oscillations in CO_2 exchange in the FL lines are not as result of instrument noise. Interestingly, as soon as the *KfCIB_FL_5B* line was transferred to LL, it did not have a large release of CO_2 of the scale observed on the previous dark to light transitions. This may possibly be linked to the change in light intensity from

400 μmoles m^{-2} s^{-1} to 100 μmol m^{-2} s^{-1} as the plants are transferred to LL. However, the large release of CO_2 was still detected when line *KfCIB1_FL _11B* was transferred into the LL conditions. (Fig. 5.11).

Once under LL conditions, *KfCIB1_FL_5B* achieved four clear oscillations in CO_2 fixation before the rhythm collapsed to arrhythmia (Fig. 5.10A, C). The free-running period of these oscillations in CO_2 fixation also shortened compared to WT, which display sustained rhythms throughout the LL time course. Similar arrythmicity also occurred in *KfCIB1_FL_11B*, but it took an extra oscillation for arrhythmia to occur (Fig. 5.11C). However, the wild type rhythm was less robust for Fig. 5.11, so it is less clear whether the arrythmicity in *FL_11B* was a major difference from the wild type. Overall, these results suggest that elevated expression of *KfCIB1* led to a more rapid collapse into CO_2 fixation arrhythmia under LL conditions, as *KfCIB1_FL_*5B was a stronger over-expresser line compared to *KfCIB1_FL_11B*.

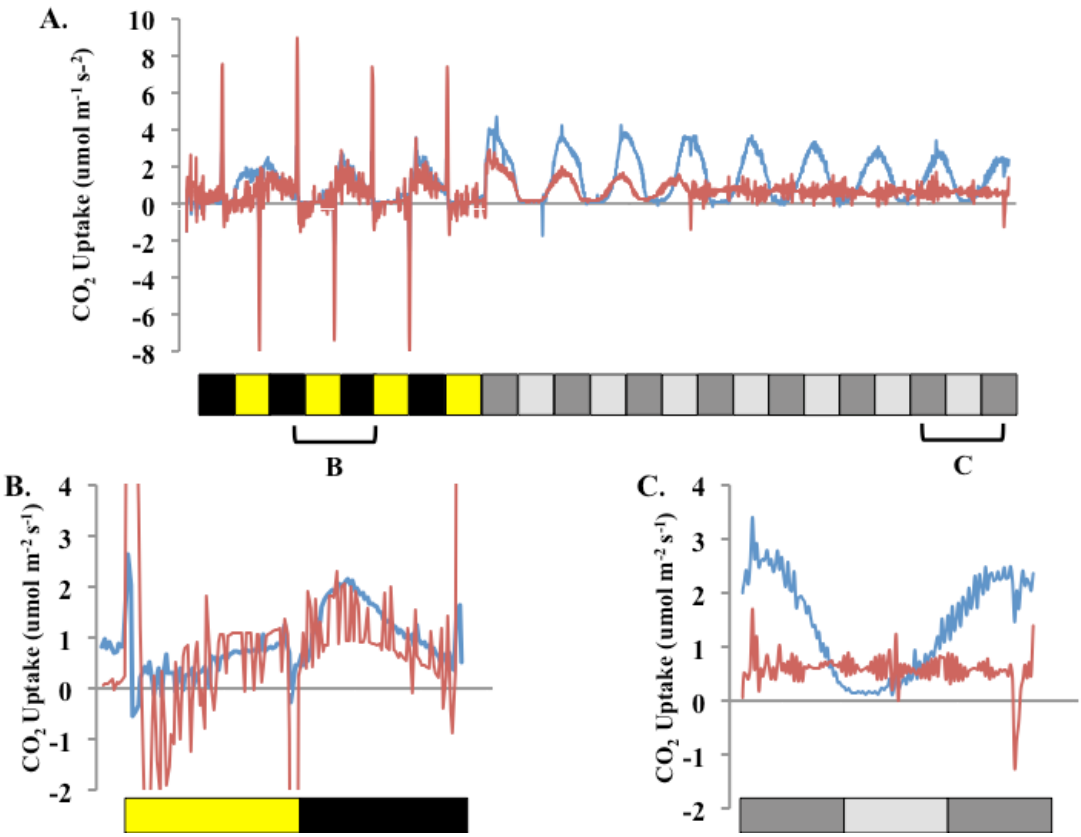

Figure 5.10 Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfCIB1_FL_5B*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfCIB1_FL_5B* is in red **A.** CO_2 fixation over 4 days LD followed by 7 days in LL. **B.** Snapshot of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C.** CO_2 fixation after entering LL, within 48h the plant is clearly already displaying a shorter period than WT and after approximately 72h becomes arrhythmic. The selected 24 h window is indicated by the C symbol below graph A.

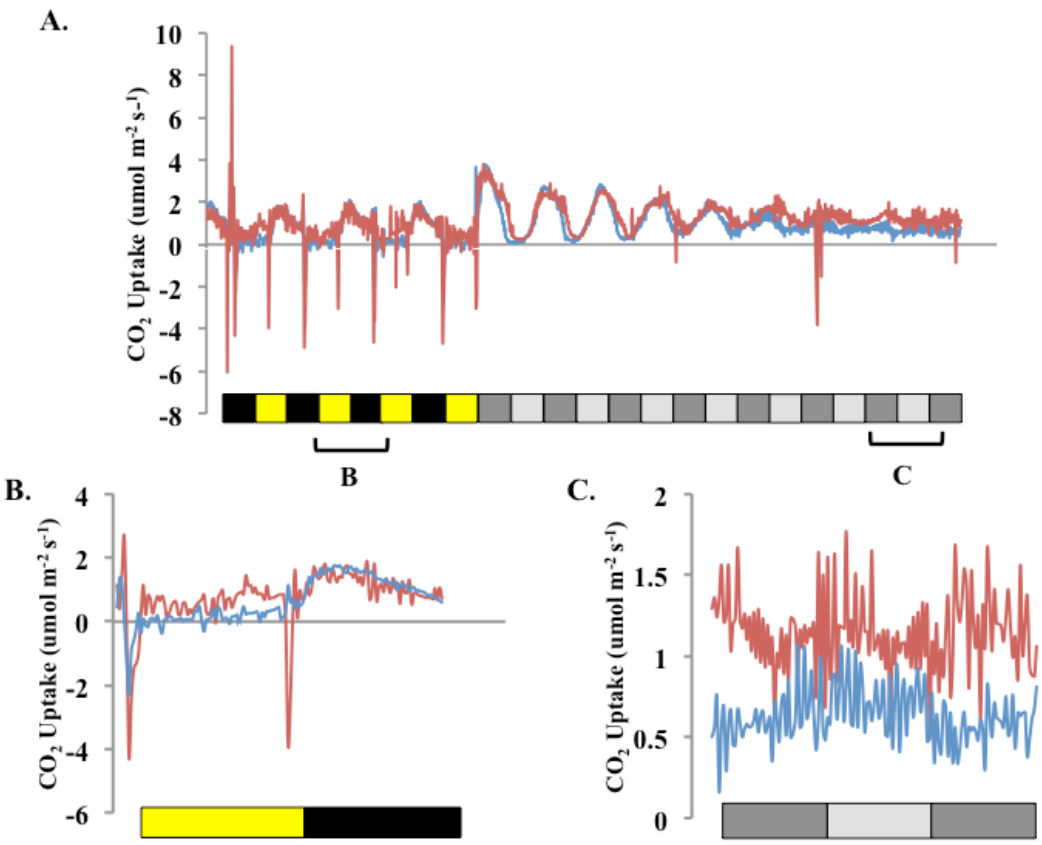

Figure 5.11 Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfCIB1_FL_11B*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfCIB1_FL_11B* is in red **A.** CO_2 fixation over 4 days LD followed by 7 days in LL. **B.** Snapshot of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C.** CO_2 fixation after entering LL. After 144h (6 days) *FL_11B* becomes arrhytmic. The selected 24 h window is indicated by the C symbol below graph A.

Using a programme designed by Ben Wareham (a PhD student in Prof. Anthony Hall's lab, University of Liverpool), the total amount of CO_2 fixed over the light and dark period could be calculated from the area under the gas exchange curve. *KfCIB_FL_5B* fixed more CO_2 than the wild type, with an 18% increase in CO_2 fixation during the light, mainly due to the greatly exaggerated phase II burst just after lights on, and the rapid oscillations between CO_2 fixation and release during the light period. 54 % less CO_2 was fixed in line *5B* during the dark period due to the rapid oscillations in CO_2 fixations (Table 5.2).

The findings were similar for *KfCIB1_FL_11B*, except *11B* fixed much more CO_2 compared to the wild type during the light (332 % more), and fixed 15 % less CO_2 during the dark period compared to the wild type (Table 5.2). This suggests that CAM, and possibly PEPC, were not working optimally in this line to fix CO_2 in the dark, as it appears that, when the stomata were opening, CO_2 was being expelled rather than fixed. It could also indicate that because strict circadian control of CAM has been altered, due to overexpression of *KfCIB1*, then futile cycling may now be occurring between key metabolic steps, such as fixing CO_2 and decarboxylating malate stores.

Table 5.2. Net amount of CO_2 fixed by each line and their associated wild types

Line	Total CO_2 fixation over the 12 h light period (µmol CO_2 fixed $m^{-2)}$	% change	Total CO_2 fixation over the 12 h dark period (µmol CO_2 fixed $m^{-2)}$	% change
WT	61912.66	-	170576.66	-
CIB FL 5B	73154.42	+18	110424.66	-54
WT	2943.87	-	51348.43	-
CIB FL 11B	12708.78	+332	43737.99	-15

5.2.4.1.2 Gas exchange measurements on overexpressing lines using the LICOR 6400XT clamp on gas exchange system

To further investigate these gas exchange results generated using the multi-channel IRGA system, in particular the remarkably large spikes of CO_2 fixation or loss from the leaf around the times that the lights changed (light to dark or dark to light), the *KfCIB1* over-expresser lines were analysed using a portable IRGA system which has a small gas exchange cuvette (2 cm^2 area) that is clamped onto the leaf whilst it is still attached to its mother plant within the growth cabinet (LICOR, LI-6400XT). This

system enabled further key physiological, photosynthesis-related, parameters to be determined, including the internal CO_2 concentration (C_i), and stomatal conductance (G_s) in addition to CO_2 fixation/ respiration. It was hoped that the use of this gas exchange system would shed further light on the CO_2 spikes measured with the multi-cuvette IRGA system. The main difference between the LICOR and the multi-channel IRGA is that the LICOR collects data from a leaf, which is still attached to a well-watered plant, and only measures a set area of the leaf. Furthermore only one plant can have measurements taken per day, meaning different lines have data collected on different days, although they are kept in the same conditions.

KfCIB1_FL_11B was entrained to 12:12 LD conditions in a Snijders Microclima cabinet, and gas exchange measurements were collected every ten minutes by clamping the leaf cuvette of the LICOR onto a LP6 leaf from each line. For *FL_11B*, there is an increase in CO_2 fixation during the light period compared to wild type (Fig. 5.12, Table 5.3) as was seen in the multi-chamber IRGA (Table 5.2). In Figure 5.12 a large increase in stomatal opening can be seen, which supports this increase in daytime CO_2 fixation. The large CO_2 spikes seen in Figures 5.11 for *FL_11B* are not seen here, but this could be because only one leaf was used for the measurement, which was attached to a well watered plant, rather than multiple LP6 leaves which had been detached from the well watered plant and had their petioles submerged in water. Despite not seeing these large spikes, stomatal conductance is clearly more sensitive to changes in light conditions, as at both light and dark transition larger fluctuations are seen in *FL_11B* than in the WT (Fig 5.12).

Furthermore, the total dark or light period CO_2 uptake calculations for *KfCIB1_FL_11B* for both IRGAs demonstrated a similar trend, as *11B* fixed more

CO_2 during the light period than WT, and less during the dark (Table 5.3). These results also showed that the results from both IRGA systems were broadly consistent with one another suggesting that each system was able to provide reliable data despite the fact that one system used detached leaves. As well as stomatal closure, leaf internal CO_2 partial pressure and CO_2 uptake were seen to drop at lights off. It was also apparent that phase IV commenced earlier and had a greater magnitude in *FL_11B* than it did in the WT. This elevated levels of phase IV was not as pronounced in the data from the CIRAS multi-cuvette gas exchange system (Fig. 5.11B). This may be because the LICOR was attached to a leaf that was part of a whole plant, whereas the other machine measured rhythms from a detached leaf placed in water, which therefore had greater potential to be under stress conditions.

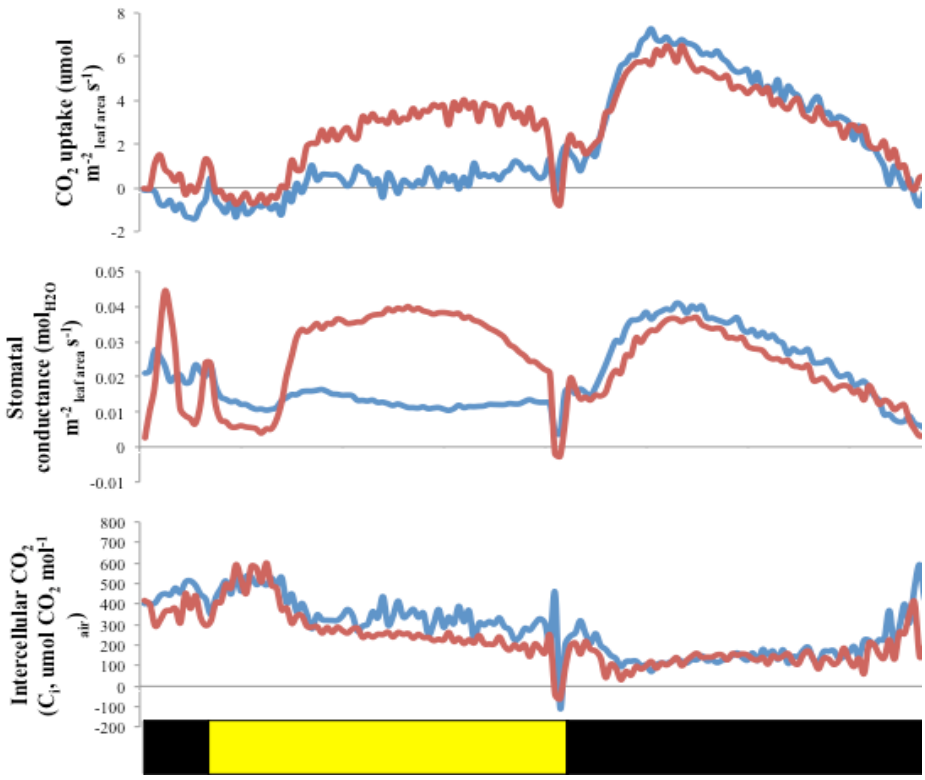

Figure 5.12. Single Leaf gas exchange data for both wild type and *KfCIB1_FL_11B* measured using the LICOR 6400XT system. Yellow bars represent lights on and black bars represent lights off. Blue line represents WT and the Red line shows *KfCIB1_FL_11B*. Three graphs show; CO_2 exchange of the leaf area in LICOR cuvette, stomatal conductance indicating how open the stomata are in that area of the leaf and the internal partial pressure of CO_2 inside the leaf. These results showed that *KfCIB1_FL_11B* fixed more CO_2 during the light period, due to fully open stomata. This may be due to more rapid drops in C_i in the light period for *KfCIB1_FL_11B* than WT.

Table 5.3. Net amount of CO_2 fixed by area of *KfCIB1_FL_11B* leaf compared to WT.

Line	Total CO_2 fixation over the 12 h light period (µmol CO_2 fixed m^{-2})	% change	Total CO_2 fixation over the 12 h dark period (µmol CO_2 fixed m^{-2})	% change
WT	822522.86	-	24282910	-
CIB FL 11B	12648948.35	93.50	22364330.03	-8.58

5.2.4.2 Impact of RNAi knockdown *KfCIB1* lines on diurnal gas exchange rhythms

The 24 h light/ dark gas exchange characteristics were also measured using the LICOR system for both *KfCIB1_RNAi* lines; *12A* and *30A*. *KfCIB1_RNAi_12A* showed higher values compared to the WT for CO_2 uptake, especially in the second half of the light, which also resulted in a drop in CO_2 concentration inside the leaf (Figs. 5.13 & 5.14; Table. 5.4**).** Lower levels of stomatal conductance can be seen throughout the light period though, which seems contradictory, if more CO_2 is being fixed. *KfCIB1_RNAi_12A* also shows changes to stomatal conductance, where at lights on and off larger fluctuations are seen than in the wild type, similar to what was seen in *FL_11B*.

KfCIB1_RNAi_30A behaved much more like the *KfCIB1_FL* over-expresser lines when measured with the LICOR system. Line *RNAi_30A* displayed a higher rate of CO_2 uptake and a much higher stomatal conductance in phase III and IV, which leads to large fluctuations in C_i, especially as lights change from on to off (Fig. 5.14). Furthermore, at the start of the measurement period, large fluctuations were detected in the transgenics, especially for stomatal conductance. This was shortly after the LICOR had been clamped onto the leaf and the door to the growth chamber had to be opened to do this, which let in a small amount of light when the plants were actually in the dark inside the chamber. This suggests that stomatal conductance in these plants was very sensitive to external changes in the environmental conditions, particularly a pulse of low light at the end of the dark period.

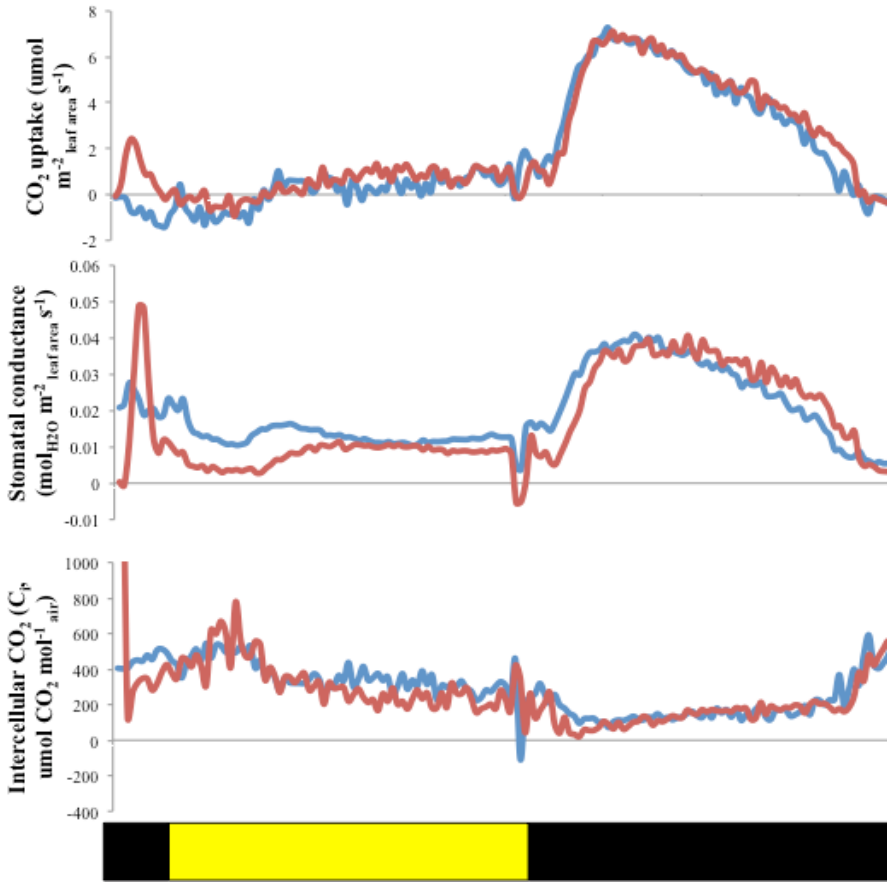

Figure 5.13. Single leaf gas exchange data for both wild type and *KfCIB1_RNAi_12A* measured with the LICOR 6400XT system. Yellow bars represent lights on and black bars represent lights off. Blue line represents WT the red line is *KfCIB1_RNAi_12A*. Three graphs show; CO_2 uptake of leaf area in the LICOR cuvette, stomatal conductance indicating how open the stomata are in that area of the leaf and intercellular CO_2, which is the pressure of CO_2 within the leaf. This shows that *KfCIB1_RNAi_12A* has very similar levels to WT, but still shows more pronounced changes in stomatal conductance and intercellular CO_2 during light transition from light to dark.

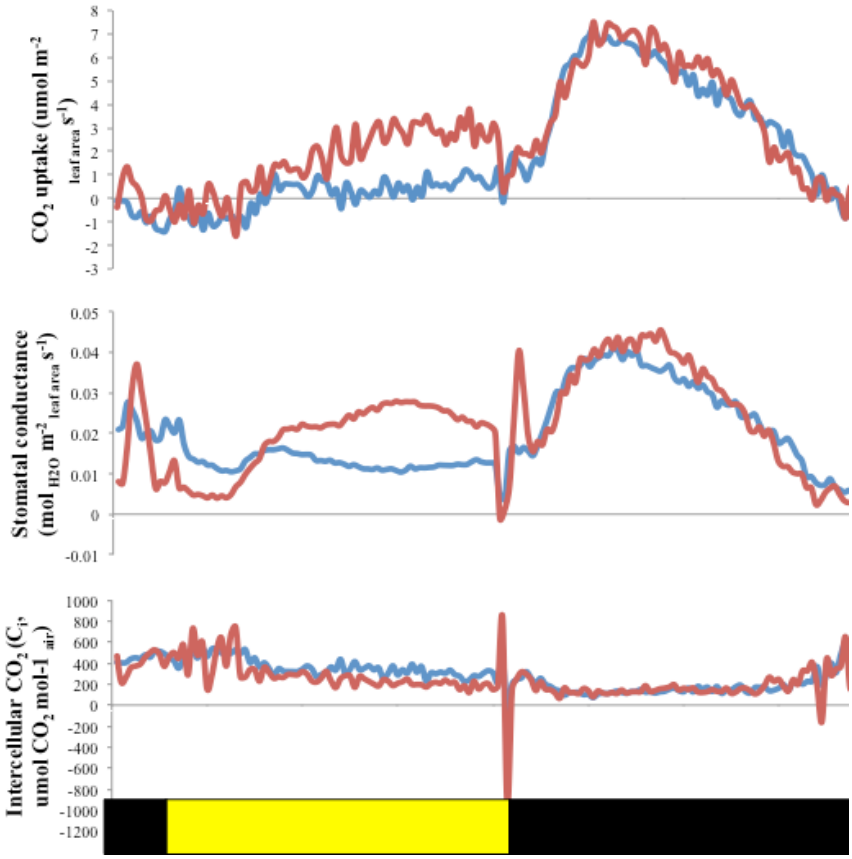

Figure 5.14. Single leaf gas exchange data for both wild type and *KfCIB1_RNAi_30A* measured with the LICOR 6400XT system. Yellow bars represent lights on and black bars represent lights off. Blue line represents WT the red line is *KfCIB1_RNAi_30A*. Three graphs show; CO_2 uptake of leaf area in the LICOR cuvette, stomatal conductance indicating how open the stomata are in that area of the leaf and intercellular CO_2, which is the pressure of CO_2 within the leaf. This shows that *KfCIB1_RNAi_30A* has very similar levels to WT, but still shows more pronounced changes in stomatal conductance and intercellular CO_2 during light transition from light to dark. More CO_2 is fixed during the light period, due to stomata being more open, which results in a lower C_i. Large fluctuations in stomatal conductance and C_i can be seen as lights change from on to off. This also results in a dip in CO_2 fixation.

Table 5.4. Net amount of CO_2 fixed by area of *KfCIB1_FL_11B* leaf compared to WT using LICOR.

Line	Total CO_2 fixation over the 12 h light period (µmol CO_2 fixed $m^{-2)}$	% change	Total CO_2 fixation over the 12 h dark period (µmol CO_2 fixed $m^{-2)}$	% change
WT	822522.86	-	24282910	-
CIB RNAi 12A	1225897.14	32.90	25214668.07	3.70
CIB RNAi 30A	9726502.64	91.54	25154344.04	3.46

5.2.5 Measurements of the end products of CAM using metabolite assays

In order to determine the impact of *KfCIB*1 transcript level changes on CAM at a biochemical level, assays were conducted to quantify the amount of malate, starch and sugars present in each of the transgenic lines,.

5.2.5.1 Impact of *KfCIB1* full-length overexpression lines on metabolite levels

5.2.5.1.1 Dark/ light fluctuations in malate concentrations in CAM leaves

In Chapter 3, basic malate screens were conducted using chlorophenol red dye to indicate pH. This assay allows quantification of malate in micromoles of malate per gram of tissue.

From the malate assays, cycling of malate clearly still occured in *KfCIB1_FL_5B* and *KfCIB1_FL_11B*. WT and both *KfCIB1_FL* lines showed a significant difference in light and dark malate concentrations (Tukey HSD, $P < 0.0001$). However, neither *FL* line displayed a significant difference relative to the WT at either the end of the light or dark period (Fig 5.15).

Despite no significant differences between wild type and transgenic lines, there are differences between the lines in malate turnover between dawn and dusk. Wild type plants had a Δ-malate of 40.1 µmol gFW^{-1}, whilst *KfCIB1_FL_5B* turned over 50.3

µmol gFW^{-1} and *KfCIB1_FL_11B* turned over 45.2 µmol gFW^{-1}, which suggests that both lines turned over more malate than wild type over the 24 h period.

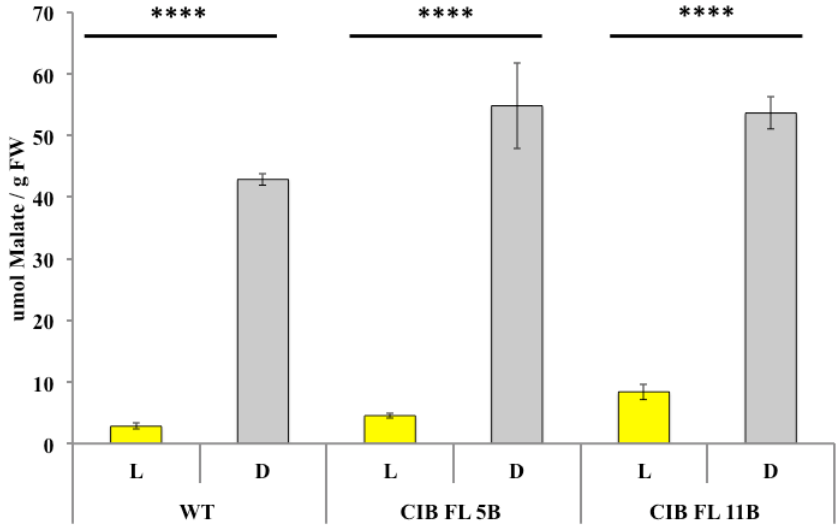

Figure 5.15. Diurnal fluctuations in malate concentrations in *KfCIB1_FL_5B* and *KfCIB1_FL_11B*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Malate levels were low at the end of the light period due to malate decarboxylation occurring during the day, and high at the end of the dark due to nocturnal CO_2 fixation resulting in malic acid accumulation. Malate levels for *KfCIB_FL* lines were higher than WT at the end of the light and dark periods, but not significantly so. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences relative to the corresponding WT value.

5.2.5.1.2 Daily fluctuations in starch

As starch and malate reciprocate in abundance between the end of the light and the end of the dark, it was important to investigate starch levels in the CAM leaves and relate these to the malate level and its daily turn over. Only visual indications in starch level were given in Chapter 3. In this section, starch levels were quantified using an enzyme linked spectrophotometric assay which allowed the calculation of starch concentration in the leaves in milligrams of starch per gram fresh weight of leaf tissue (Fig. 5.16).

There were raised levels of starch present at the end of the day in both *KfCIB1_FL* lines, yet by the end of the night, levels had fallen as low as the wild type. Wild type turned over 1 mg starch gFW^{-1} per day, whilst the over-expressing lines showed a much greater Δ-starch, with *KfCIB1_FL_5B* cycling 3.7 mg starch gFW^{-1} per day and *KfCIB1_FL_11B* turning over 4.2 mg starch gFW^{-1} per day (Fig. 5.16). Furthermore, comparing the starch concentrations at the end of the light between the wild type and *FL_5B*, the values were significantly different (Tukey HSD, P = 0.02). *KfCIB1_FL_11B* did not accumulate as much starch as *KfCIB1_FL_5B*; it accumulated an average of 4.6 mg starch by the end of the day, which dropped to an average 0.9 mg at the end of the dark (Tukey HSD, P = 0.19). There were also no large differences in the level of starch at the end of the dark, with similar levels of starch in all three lines at dawn .

This demonstrates the value of performing quatitative assays on starch in these lines, as the original screening assays with iodine staining (Chapter 3) did not detect any differences relative to the wild type at the end of dark in *KfCIB1_FL_5B,* and at either time point in line *11B*. The iodine staining also suggested that there were lower levels

of starch at the end of the light in *KfCIB1_FL_5B* compared to the wild type. This further emphasises that the value of the leaf punch iodine staining screen is limited to lines that have very large perturbations in dawn or dusk levels of starch; for example starchless mutants such as the *pgm* mutant of Arabidopsis, or starch excess mutants such as the *sex1/ gwd* mutant of Arabidopsis.

These results suggested that the FL lines were able to accumulate high levels of starch during the light period, and turn most of that starch over during the dark period (Fig. 5.16). In order to investigate further the possible destination(s) of this carbohydrate, leaf soluble sugar levels were also measured to determine whether or not the starch was being broken down to sucrose or glucose or fructose.

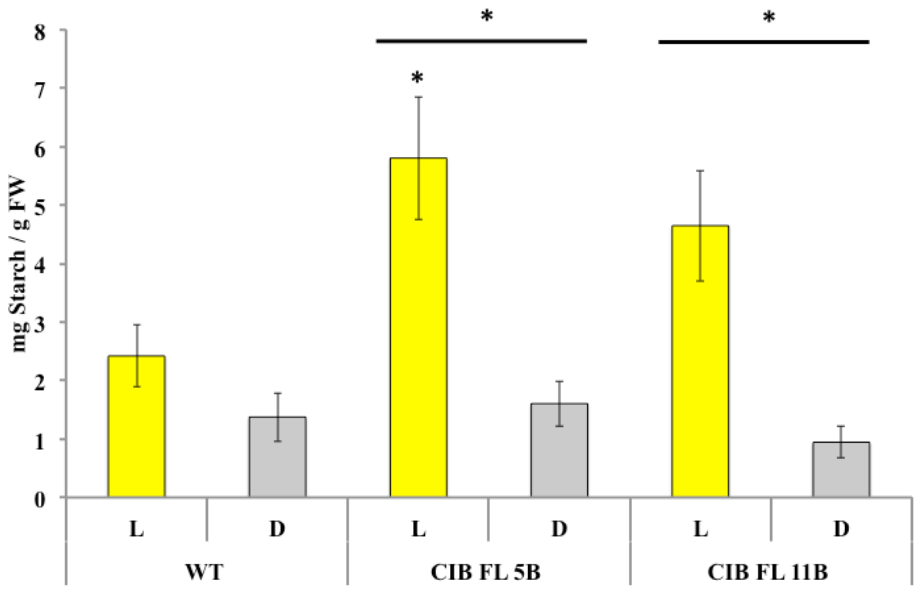

Figure 5.16. Diurnal fluctuations in starch concentrations in *KfCIB1_FL_5B* and *KfCIB1_FL_11B* relative to the wild type. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Starch levels are highest at the end of the light period, after photosynthesis has been occurring all day. Starch levels then drop to lower levels at the end of the dark, due to starch stores being mobilised to produce PEP for nocturnal CO_2 fixation. Both FL lines show an increase in accumulated starch at the end of the light period, and show no differences to wild type during the dark. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

5.2.5.1.3 Daily fluctuations in soluble sugar

Three soluble sugars were quantified for the CAM leaves from these transgenic lines; Glucose, Fructose and Sucrose.

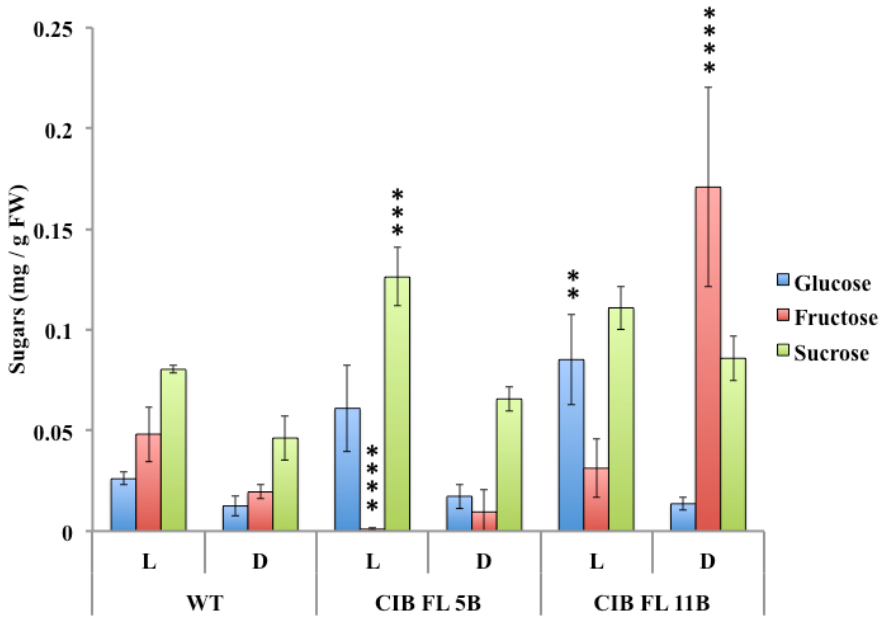

Figure 5.17. Levels of Glucose, Fructose and Sucrose in *KfCIB1* overexpressing lines, at the end of the light (L; 23:00) and dark (D; 11:00) periods. *KfCIB1_FL_5B* shows an increase in glucose levels in the light, and a reduction in fructose. *KfCIB1_FL_11B* shows an increase in glucose during the light period, but also shows a large increase in sucrose at the end of the dark period. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

For both wild type and *KfCIB1_FL_5B,* glucose levels did not significantly fluctuate between light and dark, whereas *KfCIB1_FL_11B* had a significantly higher level of glucose at the end of the light period (Fig. 5.17, Tukey HSD, P < 0.0001).

Also, the wild type had significantly higher levels of fructose than *KfCIB1_FL_5B* at the end of the light, with *5B* being close to 0 mg fructose (Tukey HSD, P < 0.0001).

However, at the end of the dark period, both lines had similar values for fructose (Fig. 5.17). *KfCIB1_FL_11B* had a large fructose peak at the end of the dark period (Fig.

5.17), which amounted to massive and highlt statistically significant increase in sucrose at the end of the dark in *FL_11B* than compared to the WT or *FL_5B* at this time (Tukey HSD, $P < 0.0001$),

Both *KfCIB1_FL* lines have increased levels of sucrose during the day, but only *5B* has significantly higher levels **(**Tukey HSD, $P = 0.004$ **).** During the dark though, *5B* and *11B* both had levels that were similar to the wild type.

Also, when looking at the total summative amounts of soluble sugars present in leaves of each line for each time point, plus a combined value for the light and dark, the *KfCIB1_FL* lines displayed increased total sugar levels at the end of the light (Fig. 5.18, Tukey HSD, $P < 0.0001$), and *KfCIB1_FL_11B* also had a large increase in dark sugars too (Tukey HSD, $P < 0.0001$). Furthermore, the total amounts of sugars present at the two time points (T), compared to wild type, showed that both lines had an overall increase in their leaf soluble sugar levels (Tukey HSD, $P < 0.0001$).

With both *KfCIB1_FL* lines the increases in light starch levels also seem to result in more sugars being present. During the dark though, *FL_5B* has similar starch and sugar values to wild type, whereas *FL_11B* shows a decrease in dark starch, which results in a significant increase in both fructose and sucrose ($P <0.0001$), suggesting that the increased daily turnover of starch may be resulting in larger pools of soluble sugars.

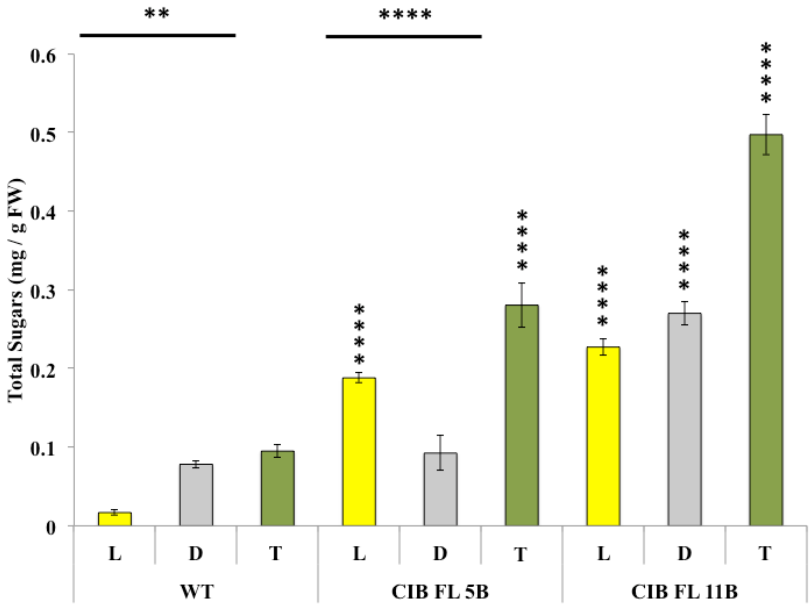

Figure 5.18 . Variation in total soluble sugar levels in WT and the *KfCIB1_FL* lines. Yellow bars indicate light sugars levels (L), grey represent dark levels (D) and green bars show the total of sugar levels (T) calculated by summing the end-of-dark and end-of-light values for all sugars. There is a clear trend that these *KfCIB1_FL* lines have higher sugar levels during the day, and total sugar turnover in a day. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

5.2.5.2 Impact of *KfCIB1* RNAi knockdown lines on metabolite levels

5.2.5.2.1 Daily fluctuations in malate

All lines showed a statistically significant difference in their daily dark/ light malate fluctuations consistent with all lines performing considerable amounts of dark CO_2 fixation via the CAM pathway (Fig. 5.19; Tukey HSD, $P < 0.0001$). Malate levels in *KfCIB1_RNAi_12A* were higher than wild type at the end of the light period (Tukey HSD, $P = 0.04$), but were very similar at the end of the dark (Fig. 5.19).

For the wild type, the Δ-malate was 40.12 µmol gFW^{-1}, whereas malate turnover for *KfCIB1_RNAi_12A* was lower at 35.62 µmol gFW^{-1}. By contrast, *KfCIB1_RNAi_30A*

showed a small increase in malate turnover compared to the wild type with a Δ-malate value of 42.52 µmol gFW^{-1}.

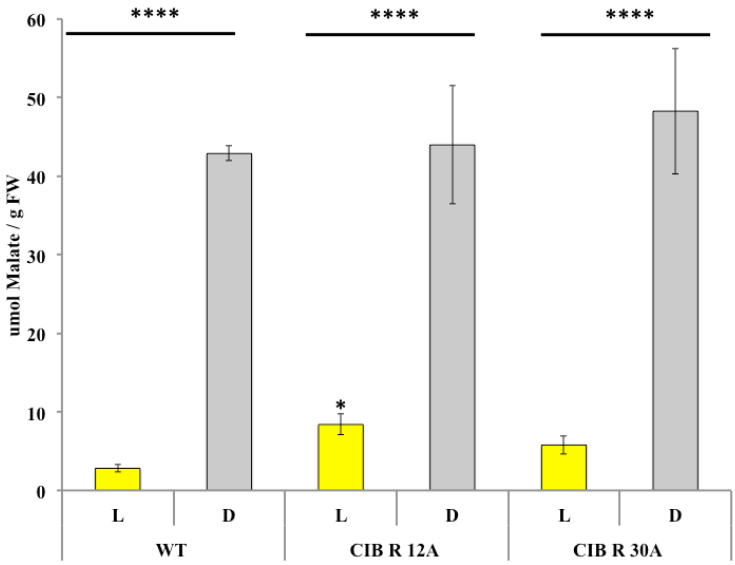

Figure 5.19. Diurnal fluctuations in malate concentrations in *KfCIB1_RNAi_12A* and *KfCIB1_RNAi_30A*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Malate levels were low at the end of the light period due to malate decarboxylation occurring during the day, and high at the end of the dark due to nocturnal CO_2 fixation resulting in malic acid accumulation. Malate levels for *KfCIB_RNAi* lines were higher than WT at the end of the light and dark periods. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences relative to the corresponding WT value.

5.2.5.2.2 Daily fluctuations in starch

As just described for malate, there were clear daily cycles of starch turnover comparing the end of light and end of dark levels of starch (WT; P = 0.01, *12A*; P = 0.05, *30A*; P = 0.02). Both *KfCIB1_RNAi* lines had on average slightly lower amounts of starch compared to the wild type at the end of the day, but during the dark period, much more starch was broken down in both RNAi knockdowns compared to the wild type (P = 0.04). Wild type average starch turnover was 1 mg starch gFW^{-1} per day. Both RNAi lines showed little variation relative to the wild type, with only a slight increase in turnover. *KfCIB1_RNAi_12A* had a Δ-starch of 1.5 mg starch gFW^{-1} per day, whilst *KfCIB1_RNAi_30A* turned over 1.7 mg starch gFW^{-1} per day (Fig. 5.20). As both of these *KfCIB1_RNAi* lines displayed an increase in the amount of starch being metabolised during the dark period, sugar assays were carried out to investigate whether the excess starch turnover was feeding into the pools of soluble sugars.

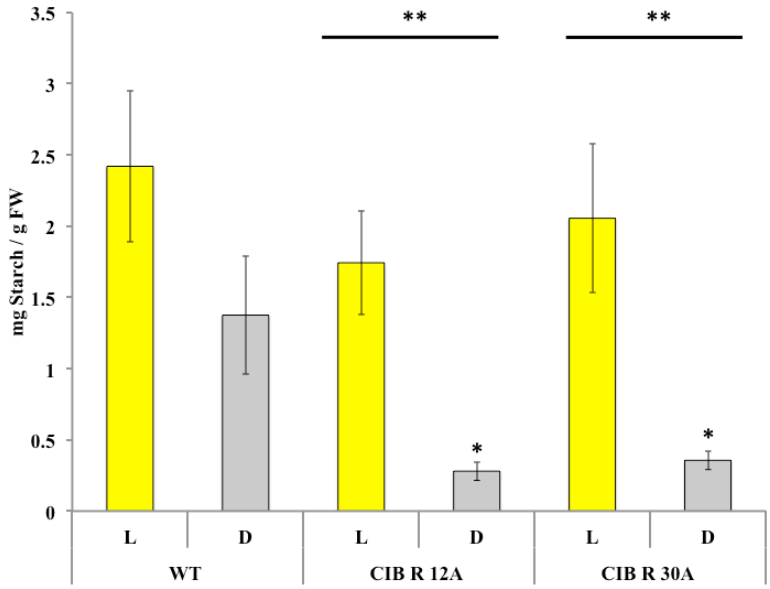

Figure 5.20. Diurnal fluctuations in starch concentrations in *KfCIB1 RNAi_12A* and *KfCIB1_RNAi_30A*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Starch levels are highest at the end of the light period, after photosynthesis has been occurring all day. Starch levels then drop to lower levels at the end of the dark, due to starch stores being mobilised to produce PEP for nocturnal CO_2 fixation. Both RNAi lines show a significant decrease in sratch levels during the dark period, and also show a reduction in average amount accumulated in the light too. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

5.2.5.2.3 Daily fluctuations in soluble sugars

Glucose, fructose and sucrose levels in the *KfCIB1_RNAi* lines were quantified and showed some large differences between these RNAi lines and WT.

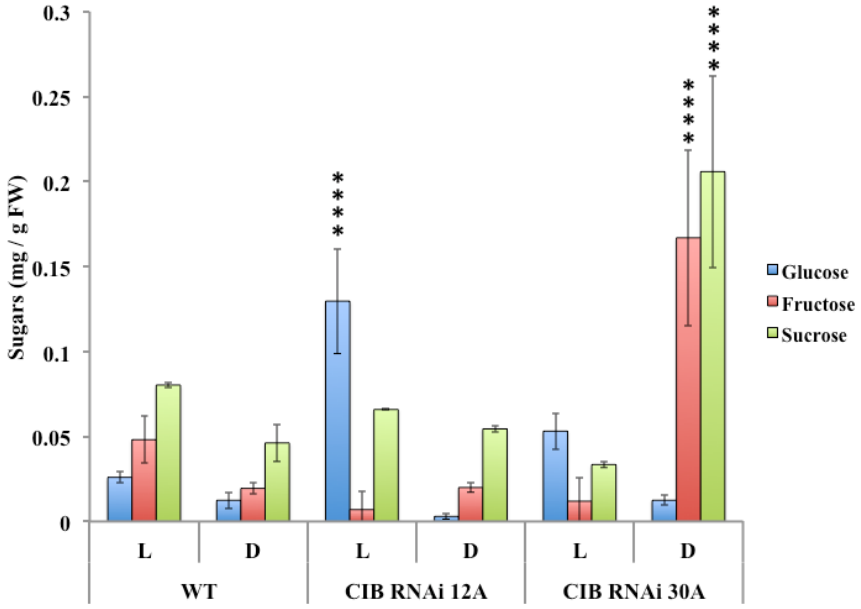

Figure 5.21. Levels of Glucose, Fructose and Sucrose in *KfCIB1_RNAi* knockdown lines, at the end of the light (L; 23:00) and dark (D; 11:00) periods. *KfCIB1_RNAi_12A* shows an increase in glucose levels in the light, and a reduction in fructose. *KfCIB1_RNAi_30A* also shows an increase in fructose at the end of the light, but also shows large increases in sucrose and fructose level at the end of the dark period. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $\leq$ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

As described above, the wild type had higher levels of sugars at the end of the light period, likely due to photosynthetic activity, and levels dropped during the dark period. WT and line *KfCIB1_RNAi_30A* lines did not display any significant fluctuations in glucose over the light/ dark cycle, whereas *KfCIB1_RNAi_12A* turned over a significant amount of glucose (Fig. 5.21, Tukey HSD, $P < 0.0001$). Line *12A* also accumulated significantly more glucose than either the WT or line *30A* at the end of the light period. The only significant fluctuations for fructose and sucrose were

seen between WT and *KfCIB1_RNAi_30A* (Tukey HSD, P < 0.0001). *KfCIB1_RNAi_30A* showed a significant increase in the level of fructose and sucrose at the end of the dark period compared to both the WT and *KfCIB1_RNAi_12A* (Tukey HSD, P < 0.0001).

Total sugar levels were also calculated for all lines. Both *KfCIB1_RNAis* accumulated more sugar than the wild type during the light period, but only the difference between WT and *KfCIB1_RNAi_12A* was significantly higher (Fig. 5.22, Tukey HSD, P = 0.04). Interestingly though, in the dark period *KfCIB1_RNAi_12A* had similar levels of sugar to the wild type, whereas *KfCIB1_RNAi_30A* had dramatically elevated levels of total soluble sugars, which was also the result obtained for the two *KfCIB1* over-expressor lines (Tukey HSD, P < 0.0001).

When comparing the total sugars present throughout the light dark cycle (T) both *KfCIB1_RNAi* lines showed significant increases compared to the wild type (Tukey HSD, 12A; P = 0.04, 30A; P < 0.0001).

For *RNAi_30A*, the significant decrease in starch level, leads to a significant increase in fructose and sucrose in the leaf, suggesting fewer sugars are being converted to starch, or there is futile cycling between the two. With *RNAi_12A*, which shows the largest transcript reduction, there is both lower starch and lower sugars, suggesting futile cycling of metabolites may be occurring due to perturbation of the circadian control of CAM.

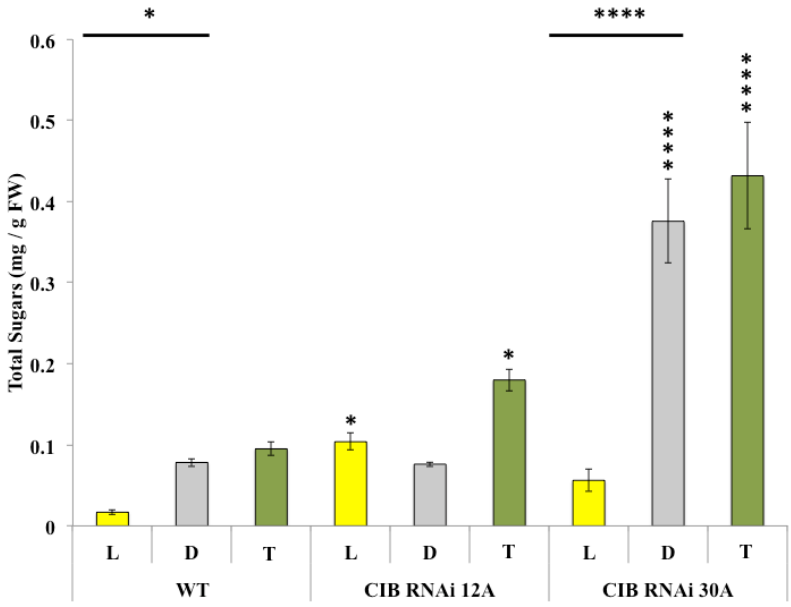

Figure 5.22. Variation in total soluble sugar levels in WT and the *KfCIB1_RNAi* lines. Yellow bars indicate light sugars levels (L), grey represent dark levels (D) and green bars show the total of sugar levels (T) calculated by summing the end-of-dark and end-of-light values for all sugars. There is a clear trend that these *KfCIB1_RNAi* lines have higher total sugar turnover in a day. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

5.2.6 Immunoblot determination of the relative protein abundance of key enzymes associated with CAM

Protein levels of key CAM enzymes were investigated to determine whether the over-expression and RNAi knockdown of *KfCIB1* had a knock on effect on the abundance and regulation of key CAM enzymes.

5.2.6.1 Protein abundance in *KfCIB1* full-length overexpression lines

Due to issues with extraction of proteins and also time constraints with *KfCIB1_FL_11B,* it was not possible to complete immunoblot analysis on this line. Therefore, only the immunblot results for *KfCIB1_FL_5B* are described in this section.

For *KfCIB_FL_5B* no differences in protein abundance relative to WT were detected for PEPC protein levels (Fig. 5.23A). A difference was however detected for PEPC phosphorylation state in the dark (Fig. 5.23B). WT showed a very low amount of PEPC phosphorylation at the start of the dark, which then increased, peaking in the middle of the dark before declining slightly in the 2 h before dawn sample. Phospho-PEPC could only be detected in *KfCIB1_FL_5B* at 6 and 10 h into the 12 h dark period, but at both time points the phospho-PEPC signal was lower than the corresponding wild type (Fig. 5.23B)

For NAD-ME α (Fig. 5.23C) and β (Fig. 5.23D), both showed slightly increased levels of protein compared to wild type at the end of the dark and start of the light period, which fits with the transcript data shown in Figure 5.4.

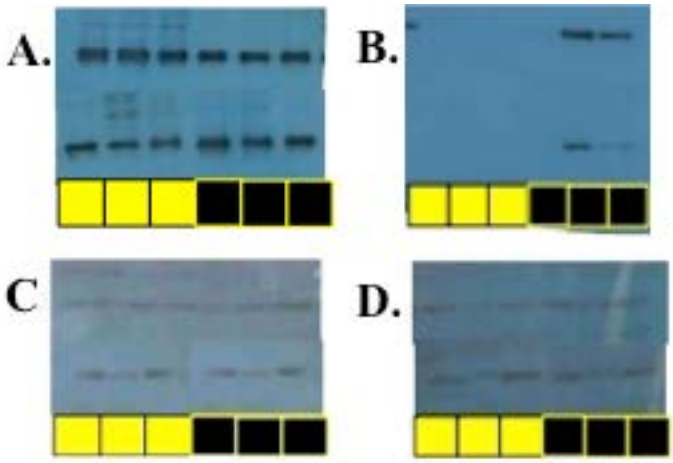

Figure 5.23. Immunoblot analysis revealed that several key CAM proteins in the *KfCIB1_FL_5B* over-expressor line. **A.** Immunoblot for PEPC. **B.** Immunoblot for phosphorylated PEPC; **C.** Immunoblot for NADME alpha, **D.** Immunoblot for NADME beta. (Top line –WT, bottom line – *KfCIB1_FL_5B*).

5.2.6.2 Protein abundance in *KfCIB1* RNAi Knockdowns

The first proteins screened for were PEPC and phosphorylated PEPC, as although the changes in transcript level were small, there were still slight increases during the dark period (Figs. 5.6 & 5.7).

WT, *KfCIB1_RNAi_12A* and *KfCIB1_RNAi_30A* all showed similar levels of PEPC throughout the light/ dark cycle, confirming equal loading of protein in each well of the SDS-PAGE gel. When looking at phospho-PEPC levels, wild type and *KfCIB1_RNAi_12A* both showed very similar phosphorylation patterns in the dark, with high phosphorylation at 6 h and 10 h into the 12 h dark period (Fig 5.24B). Interestingly though, *KfCIB1_RNAi_30A* had a drop in phosphorylation by 10 h dark, indicating that less of the PEPC was phosphorylated in the leaves of this line at the end of the dark period (Fig. 5.24B). This explains why this line shows a quicker decline in dark CO_2 fixation, although the decline is not a rapid as may be expected from having barely any phosphorylated PEPC being present.

Blots were also carried out for NAD-ME α and β. For wild type and *KfCIB1_RNAi_30A,* the protein levels were very similar, with constant levels throughout the 24 h period (Fig. 5.24A;B). However, *KfCIB1_RNAi_12A* displayed lower levels of NAD-ME α throughout the dark period (Fig 5.24C).

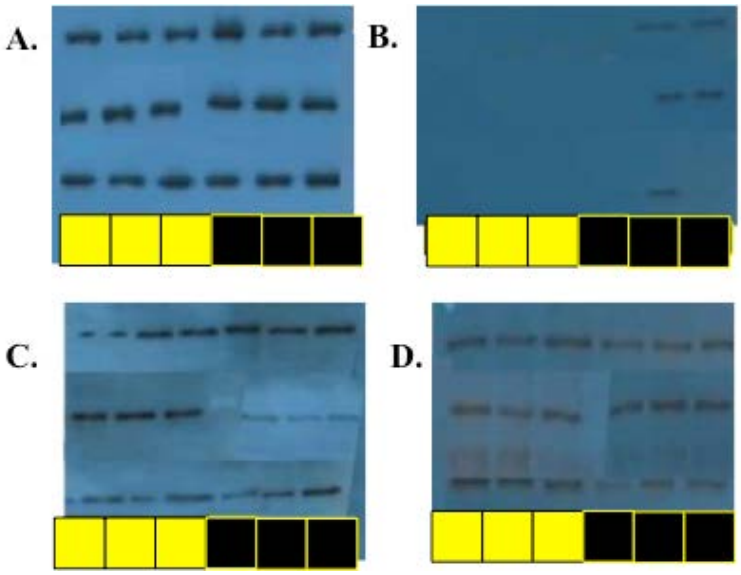

Figure 5.24. Immunoblotting revealed that several key CAM proteins in the *KfCIB1_RNAi* knockdown lines. **A.** Immunoblot for PEPC. **B.** Immunoblot for phosphorylated PEPC; **C.** Immunoblot for NADME alpha, **D.** Immunoblot for NADME beta. (Top line –WT, middle line – *KfCIB1_RNAi_12A* and bottom line – *KfCIB1_RNAi_30A*).

5.2.7 Enzyme Assay conducted on transgenic *KfCIB1* lines

As well as using immunoblotting to detect the presence of the protein, it was also important to determine their activity. PEPC was assayed for total extractable activity and the apparent K_i of PEPC for malate was assayed using desalted extracts. Total extractable PEPC activity was expressed both in μmol min^{-1} mg $protein^{-1}$ and per mg chlorophyll. Samples were collected in middle of the dark (6 hours into the 12 h dark period).

The percentage inhibition of PEPC in rapidly desalted extracts by different concentrations of malate was calculated for both light (6 h into 12 h light period) and dark (6h into 12 h dark period) samples, to investigate the *in planta* differences in

PEPC phosphorylation state, and therefore how much CO_2 could be fixed due to malate inhibition of PEPC.

5.2.7.1 Enzyme assay results for *KfCIB1* full-length overexpression lines

5.2.7.1.1 Total extractable activity of PEPC

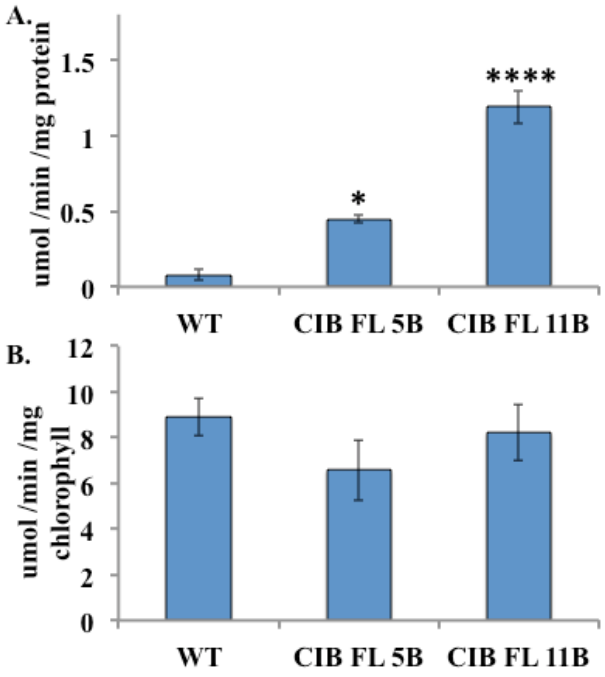

Figure 5.25. Total extractable activity of PEPC during the dark period (6h dark) for WT and 2 *KfCIB1_FL* lines. **A.** Total extractable activity of PEPC per mg of protein. **B.** Total extractable activity of PEPC per mg of Chlorophyll. Both KfCIB1 lines show vastly increased levels in activity compared to the wild type when expressed per mg protein, but when expressed per mg chlorophyll both show a small reduction in activity. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $\leq$ 0.0001. Stars above one column represent significant differences to the corresponding WT value.

For both *KfCIB1_FL* lines, activity was significantly higher than WT when expressed per mg protein. *KfCIB1_FL_5B* showed a 2.9-fold increase in activity (One-Way ANOVA, $P = 0.02$), whilst *KfCIB1_FL_11B* displayed a 9.6-fold increase (Fig. 5.25; One-Way ANOVA, $P < 0.0001$). There was also a significant difference between the *KfCIB1_FL* lines, showing that *11B*, which has a lower fold-change in *KfCIB1* gene

expression, actually had a larger increase in PEPC activity (One-Way ANOVA, P = 0.0004).

PEPC activity was also expressed per mg chlorophyll. This resulted in neither line showing a significant difference to wild type, and on average, both lines showed a decrease in activity, which fits better with the decrease in PEPC phosphorylation dectected in the phospho-PEPC immunoblot for *KfCIB1_FL_5B* and also the decreases seen in dark CO_2 fixation (Fig. 5.23A).

5.2.7.1.2 Measurement of malate sensitivity of PEPC

For the wild type, the K_i during the light was 1.9 mM malate, whereas in the dark the K_i rose to 3.7 mM malate, this confirmed the expected light/ dark variation in PEPC K_i for malate in the wild type (Table 5.5). For *KfCIB1_FL_5B,* the K_i was much lower than the wild type at both time points, but did increase during the dark period. However, the dark K_i for *FL_5B* did not reach the light K_i level of the WT (Table 5.5). For *KfCIB1_FL_11B,* the K_i was still lower than wild type at each time point, but was not as low as the values determined for *FL_5B*. In all of these lines though, the light K_i approximately doubled from light to dark.

Table 5.5. Variation in the apparent K_i of PEPC for malate in rapidly desalted extracts of CAM leaves of WT and the *KFCIB1_FL* transgenic lines sampled at the middle of the light (6h L) and the middle of the dark (6h D).

Line	Light Malate K_i (mM)	Dark Malate K_i (mM)
WT	1.9	3.7
KfCIB1 FL 5B	0.6	1.3
KfCIB1 FL 11B	1.2	2.4

5.2.7.2 Enzyme assay results for *KfCIB1* RNAi knockdown lines

5.2.7.2.1 Total extractable activity of PEPC

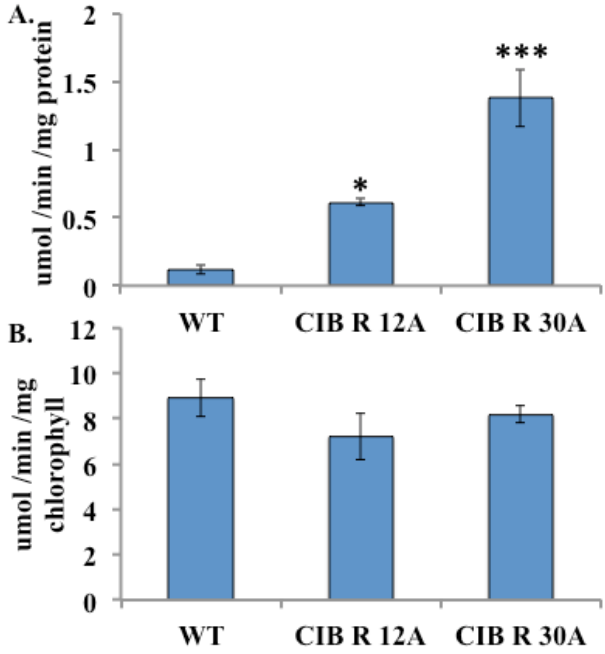

Figure 5.26. Total extractable activity of PEPC during the dark period (6h dark) for WT and 2 *KfCIB1_RNAi* lines. A. Total extractable activity of PEPC per mg of protein. B. Total extractable activity of PEPC per mg of Chlorophyll. Both KfCIB1 lines show vastly increased levels in activity compared to the wild type when expressed per mg protein, but when expressed per mg chlorophyll both show a small reduction in activity. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars above one column represent significant differences to the corresponding WT value.

When activity was expressed per mg of protein, both *KfCIB1_RNAi* lines displayed increases in PEPC activity compared to the wild type, with *KfCIB1_RNAi_12A* showing a 4-fold increase (One-Way ANOVA, P = 0.06), and line *KfCIB1_RNAi_30A* showing a 10.3 fold increase (Fig. 5.26; One-Way ANOVA, P = 0.0008).

As with the *KfCIB1_FL* lines, both *KfCIB1_RNAi* lines also showed an average decrease in activity, and no significant differences to wild type, when the total extractable activity was expressed per mg chlorophyll, suggesting that it may just be

the level of total protein in these transgenic lines that is perturbed, rather than there being changes in PEPC activity.

5.2.7.2.2 Measurement of malate sensitivity of PEPC

Wild type had a K_i of 1.9 mM malate at the start of the light period and 3.7mM in middle of the dark. Both *KfCIB1_RNAi* lines showed a reduction in K_i relative to this. *KfCIB1_RNAi_12A* showed a light K_i of 1.2, and in the dark this almost doubled, reaching 2.3 mM (Table 5.6). For *KfCIB1_RNAi_30A*, the K_i displayed a large reduction, with the dark K_i not even reaching the value seen in wild type light samples. This was the only *KfCIB1* transgenic line that did not manage to approximately double its K_i between light and dark conditions. Interestingly, this line had the highest transcript reduction in the level of the target gene, *KfCIB1*.

Table 5.6. Variation in the apparent K_i of PEPC for malate in rapidly desalted extracts of CAM leaves of WT and the *KfCIB1_RNAi* transgenic lines sampled in the middle of the light (6 h L) and the middle of the dark (6 h D).

Line	Light K_i (mM malate)	Dark K_i (mM malate)
WT	1.9	3.7
KfCIB1 RNAi 12A	1.2	2.3
KfCIB1 RNAi 30A	0.8	1.1

5.2.8 The impact of drought stress on the growth and stress-related physiology of the transgenic lines

Lines were all grown from developmentally-synchronised adventitious leaf margin plantlets, and were raised in greenhouse growth conditions **(**Materials and Methods 2.11.1) for 6 months. Soil moisture readings were collected to quantify the water content of the soil (Appendix: Figure S5.1; Table S5.1), and then approximately half of the population of plants for each line was transferred to drought conditions (water was with-held), whilst the other half remained in well-watered. After a further 31-days of being exposed to these different growth conditions, plants were harvested and soil moisture content was re-measured (Appendix S5.1). Photographs of each line are in Figures 5.28 for FL lines and 5.32 for RNAi lines.

5.2.8.1 Drought effects on *KfCIB1* Full-Length overexpression lines

5.2.8.1.1 Above-ground biomass

For line *KfCIB1_FL_5B,* 8 plants were subjected to each condition and for *KfCIB1_FL_11B* there were 11 of each. All lines showed large differences in fresh and dry weight and water content between the two growth conditions (Two-Way ANOVA, $P < 0.0001$).

All lines appeared to perform approximately equally in well-watered conditions, and also in drought conditions. (Well watered; WT = 371 g, *5B* = 393 g, *11B*; 316 g, Drought; WT = 76 g, *5B* = 76 g, *11B* = 96 g; Fig. 5.27A, C &E).

The dry weights showed slightly larger differences between well-watered and drought-stressed conditions than the wet weights, with both *KfCIB1_FL*s performing slightly better than the wild type in terms of dry weight productivity of above-ground tissues under the drought-stress conditions. *KfCIB1_FL_11B* showed no statistical

difference between well-watered and drought-stressed dry weights of its above ground tissues (Well Watered; WT = 18 g, *5B* = 20 g, *11B* = 16 g, Drought; WT = 10 g, *5B* = 12 g, *11B* = 12 g). Wild type showed the largest difference in weight between being drought-stressed and well-watered suggesting this was more strongly affected by drought than the *KfCIB1* over-expressers (Two-Way ANOVA, $P < 0.0001$).

When looking at moisture content in the plants, in well-watered conditions there was no difference, all lines has a high water content of around 85-95%. Under drought, the lines showed differences, although none were statistically significant.

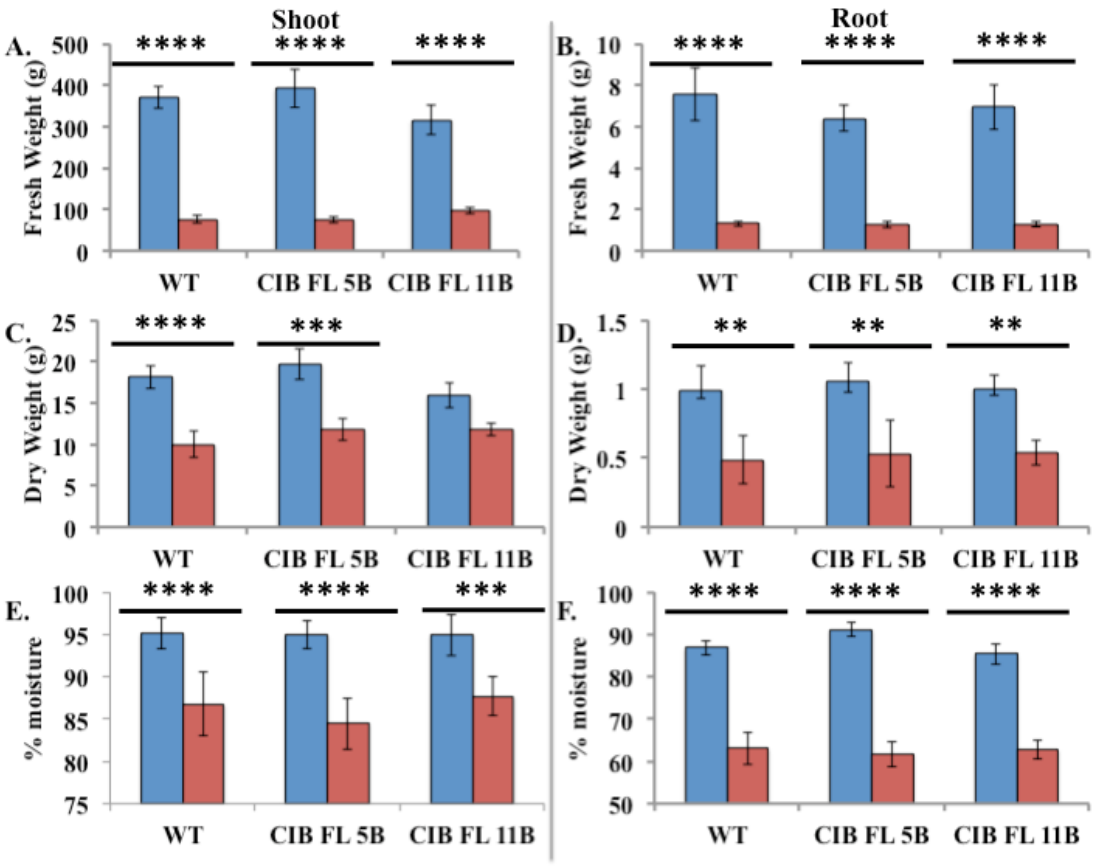

Figure 5.27. Fresh and dry weight and relative moisture content for well-watered and drought-stressed WT and FL lines. Blue bars show well watered lines, and red bars show drought-stressed lines. A, C & E show data collected for above ground weight, and B, D & F show data for below ground tissues (roots). A and B show fresh weight of samples at time of harvest, C and D, show dry weight after tissues were placed in a drying oven at 70˚C until a constant weight was reached. E and F show the % moisture held within the plant; calculated from the fresh and dry weights. A Two-Way ANOVA Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one column, e.g. well-watered/ drought represent significant differences compared to the corresponding WT value.

5.2.8.1.2. Below-ground root biomass

As with the above ground material, each line showed significant differences between wet and dry weights comparing well-watered and drought-stressed conditions (Fig. 5.27B, D & F; Tukey HSD, $P < 0.0001$).

The weight of the roots were also recorded, but these showed no significant differences in fresh weight between wild type and the two *KfCIB1_FLs* in well-

watered, and also when in drought conditions (Fresh Weight: Well watered; WT = 8 g, *5B* = 6 g, *11B* = 7 g; Drought = WT, *5B*, *11B* all = 1.3 g; Tukey HSD, P = 0.09 – 0.9).

Root dry weights also showed no significant differences between the lines (Dry Weight ; Well Watered; WT, *5B, 11B* all = 1 g, Drought = WT, *5B, 11B* all = 0.5g; Tukey HSD, P = 0.8 - >0.999).

When looking at moisture levels in the roots of these plants, there were large differences between well-watered and drought-stress conditions as expected, (Tukey HSD, WT; P = 0.004, 5B; P = 0.002, 11B; P = 0.04), but there were no differences between the transgenic lines and the wild type, which all performed approximately equally well (Well Watered; WT = 87 %, *5B* = 84 %, *11B* = 86 %, Drought; WT = 63 %, *5B* = 58 %, *11B* = 58 %).

5.2.8.1.3 Induction of UV protective anthocyanin pigmentation in response to drought-stress

Simply from looking at the three lines, WT and the two *KfCIB1_FLs*, it was clear that the pigment levels in the plants changed in response to drought-stress. In all three, purple anthocyanin pigments could be seen, which are produced as part of the plant's response to drought-stress. The WT had small amounts of anthocyanins in its leaves in well-watered conditions, whereas no anthocyanin was detected for the *KfCIB1_FLs* (WT = 0.001 μM cm^{-2}, 5B = <0.0001 μM cm^{-2}, 11B = <0.0001 μM cm^{-2}; Fig. 5.28; Tukey HSD, P < 0.0001). Following the drought-stress treatment, WT leaves had greatly elevated levels of anthocyanin pigments present, whereas much lower amounts were detected in the *KfCIB1_FL* lines (WT = 0.008 μM cm^{-2}, 5B = 0.002 μM cm^{-2}, 11B = 0.004 μM cm^{-2}). This could be due to LP6 being sampled for these

measurements. On the *KfCIB1_FL* lines, only the first few sets of leaves went a dark purple, whereas nearly all leaves changed colour in the WT. This therefore suggests that *KfCIB1* over-expression may lead to a decrease in anthocyanin production in response to drought-stress, at least in mature leaves.

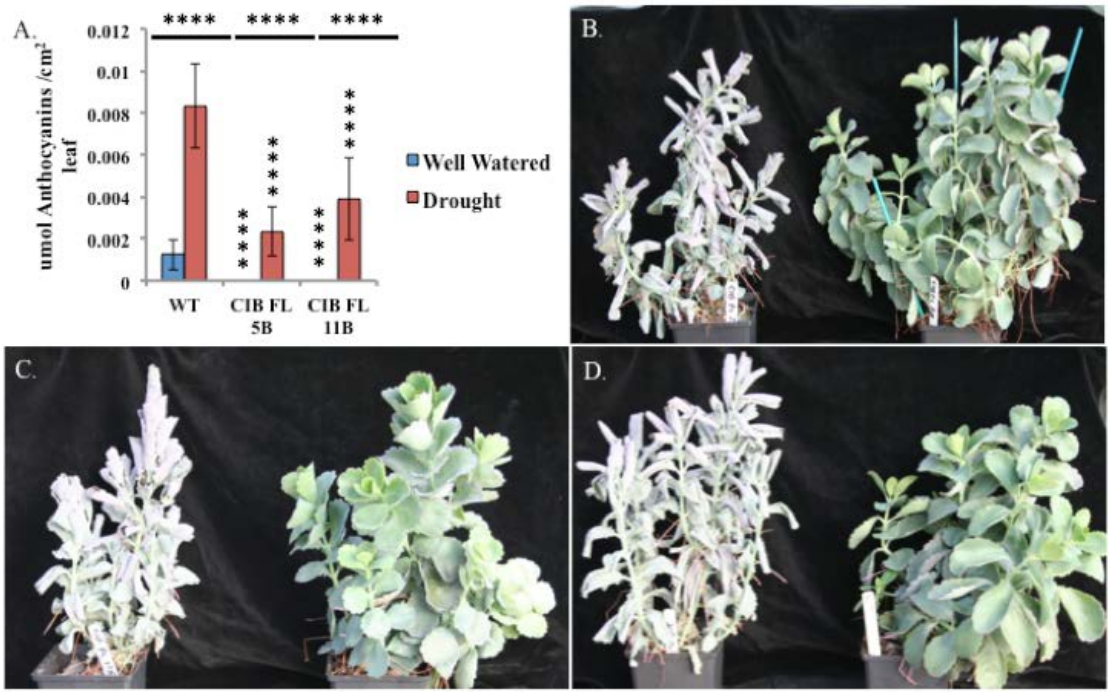

Figure 5.28. Variation in the level of anthocyanin leaf pigments in response to drought-stress for WT and *KfCIB1_FL_5B* and *11B*. **A.** Quantified levels of Anthocyanins in WT and the *KfCIB1_FL* lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one column, e.g. drought, represent significant differences to the corresponding WT value. **B.** Images of drought-stressed (Left) and well-watered (Right) WT. **C.** Images of drought-stressed (Left) and well-watered (Right) *KfCIB1_FL_5B*. **D.** Images of drought-stressed (Left) and well-watered (Right) *KfCIB1_FL_11B.*

Interestingly in all lines there was a significant increase in all of the measured photosynthetic pigments in response drought, in addition to the increase in anthocyanins. This included chlorophyll A and B, and the carotenoids (Fig. 5.29; Tukey HSD, P < 0.0001). The smallest increase in chlorophyll and carotenoid levels between well-watered and drought-stressed conditions was detected in the wild type, whereas the wild type showed the largest pigment change for the anthocyanins. When

comparing lines in drought or in well-watered conditions though, there was little if any difference for any of the measured photosynthetic pigments.

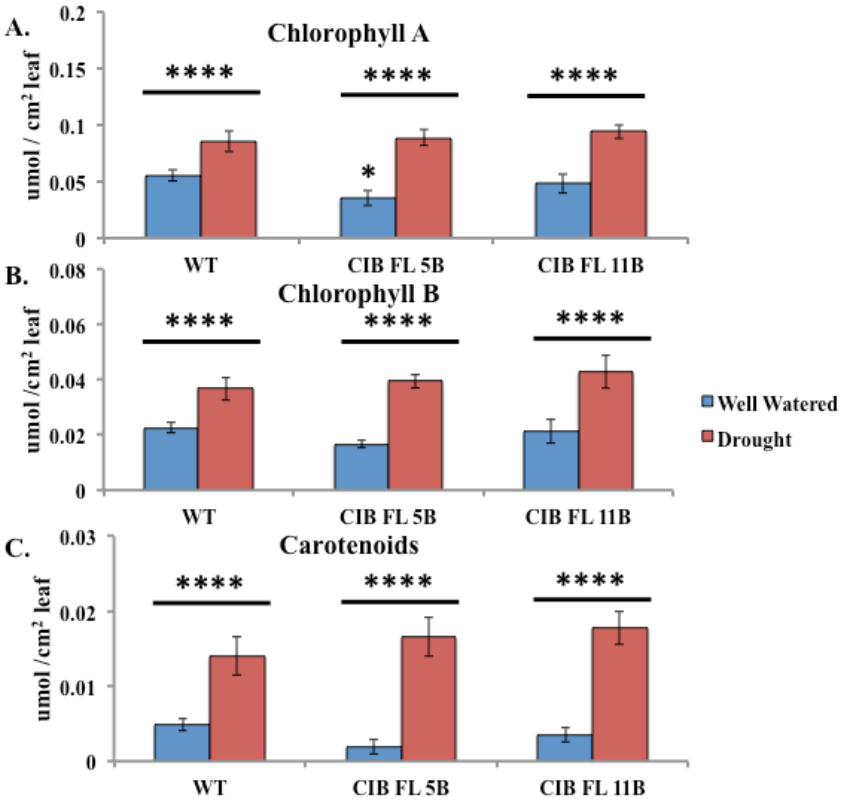

Figure 5.29. Variation in the levels of photosynthetic pigment in response to drought-stress in WT and *KfCIB1_FL* lines. **A.** Changes in the levels of Chlorophyll A in response to drought. **B.** Changes in the levels of Chlorophyll B in response to drought. **C.** Changes in the levels of Carotenoids in response to drought. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; **** shows P = ≤ 0.0001. Stars over bars at the top of the graph represent lines that are significantly different between treatments.

5.2.8.1.4 CAM-associated metabolite levels under drought-stressed conditions

Levels of malate were measured at dawn and dusk in order to determine the level of CAM occurring in the drought-stressed plants. In well-watered conditions, significant differences were detected between light and dark levels in each line, indicating CAM was operating efficiently in all lines (Fig 5.30A; $P < 0.0001$). Under the drought-stress treatment. *KfCIB1_FL_5B* no longer showed light/ dark fluctuations in malate levels. WT and *KfCIB1_FL_11B*, still show significant differences in light and dark

malate concentrations (Fig. 5.30B; P < 0.0001). Furthermore, when comparing well-watered and drought-stressed end of dark malate levels, neither wild type (P > 0.999) nor *KfCIB1_FL_11B* (P = 0.33), showed a significant difference between the two conditions. *KfCIB1_ FL_5B* did show a significant difference, suggesting *FL_5B* may have transitioned to CAM-idling due to the drought. (P < 0.0001).

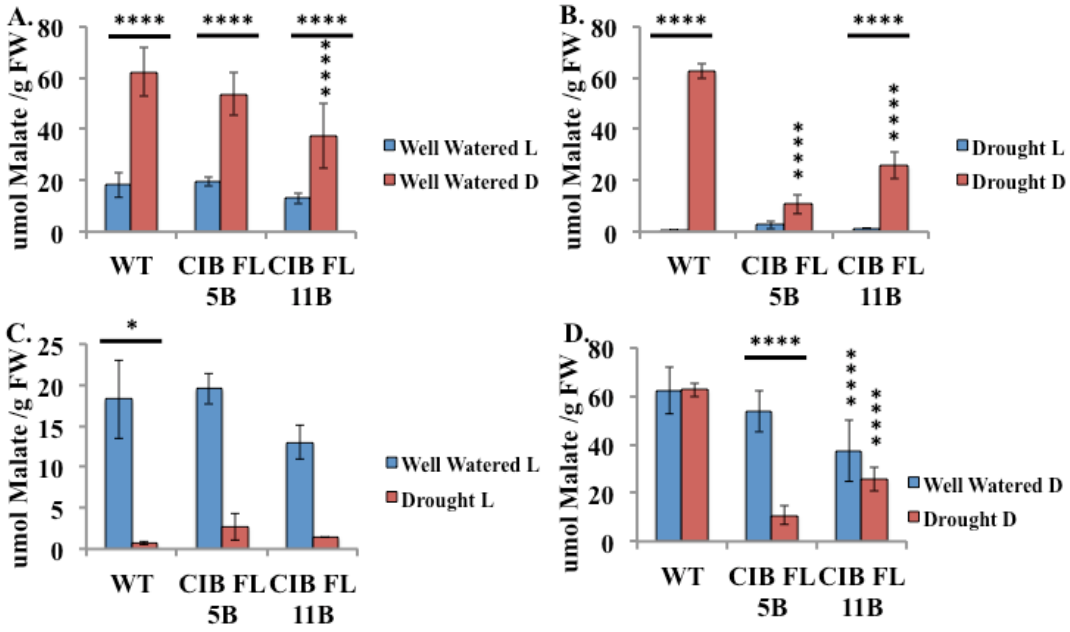

Figure 5.30. Daily fluctuations in malate content of CAM leaves of the WT and FL overexpresser lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. well-watered and Drought, and stars above one column, e.g. drought, represent significant differences from the corresponding WT value.

5.2.8.2 Drought effects of *KfCIB1* RNAi knockdown lines

5.2.8.2.1 Above-ground biomass

For *KfCIB1_RNAi_12A*, 6 clonal plants were placed into drought-stressed and well-watered conditions, and for *KfCIB1_RNAi_30A* 7 plants were grown under each condition. All lines showed large differences between well-watered and drought-stressed conditions after 31 days (Fig. 5.31A, C & E; P < 0.0001).

For both *KfCIB1_RNAi* lines, their weight at harvest after 31days was very similar to that of WT, for both well-watered and drought-stressed conditions (Well Watered; WT = 371 g, *12A* = 396 g, *30A* = 318 g; Drought; WT, *12A, 30A* all averaged 76 g). When the harvested tissues from these lines were dried down, to remove all moisture, WT achieved significantly higher dry weight yield in well-watered conditions compared to drought (Well Watered WT = 18 g, Drought WT = 10 g; Tukey HSD, $P < 0.0001$). Both *KfCIB1_RNAi* lines showed no statistical differences between well-watered and drought, suggesting lines grew similarly in both conditions, especially for *KfCIB1_RNAi_12A* (Well Watered; *12A* = 16 g, *30A* = 15 g; Drought; *12A* = 15 g, *30A* = 10 g).

All lines showed large differences between moisture levels in the two growth conditions (Tukey HSD, WT; $P < 0.0001$, *12A*; $P = 0.05$, *30A*; $P < 0.0001$). *KfCIB1_RNAi_12A* had significantly reduced percentage moisture compared to wild type in drought-stressed conditions, whilst in well-watered conditions all three lines had an approximately equal water content in their tissues (Well watered; WT = 95 %, 12A = 96 %, 30A = 95 %; Drought; WT = 87 %, 12A = 81 %, 30A = 87 %).

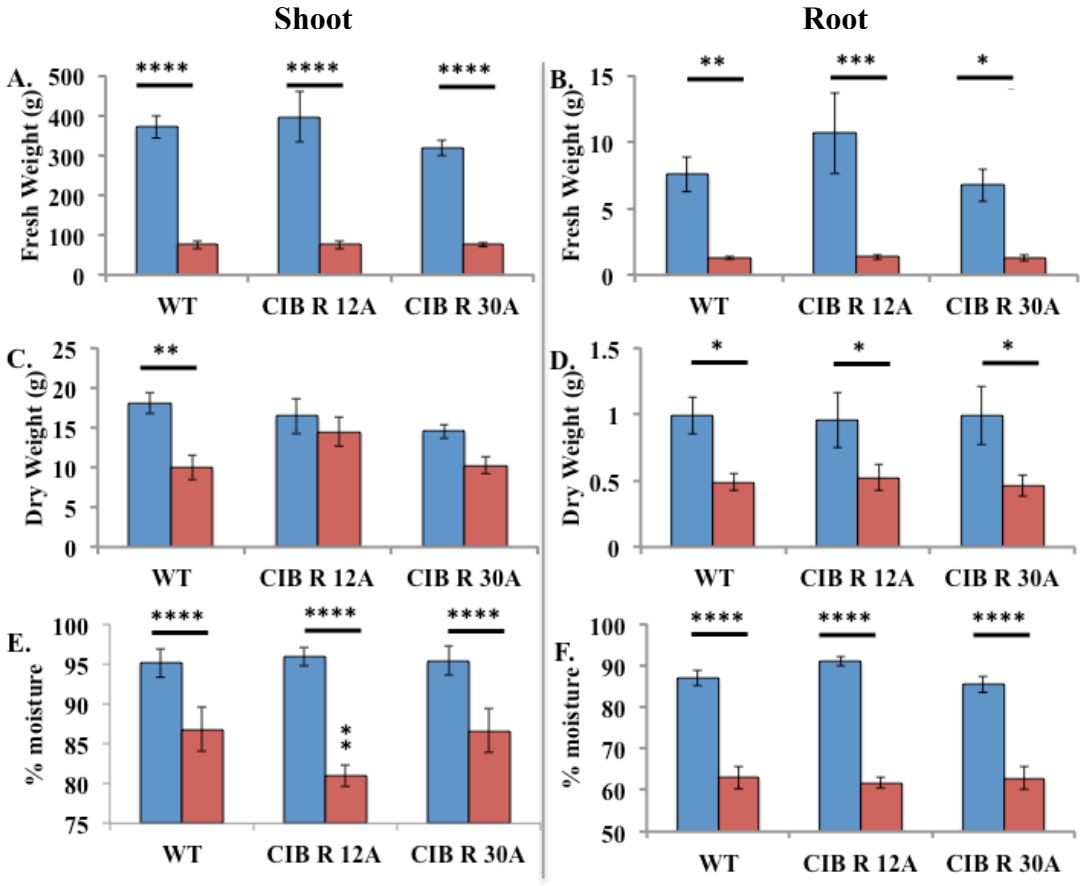

Figure 5.31. Fresh and dry weight and relative moisture content for well-watered and drought-stressed WT and RNAi lines. Blue bars show well-watered lines, and red bars show drought-stressed lines. A, C & E show data collected for above ground weight, and B, D & F show data for below ground tissues (roots). A and B show fresh weight of samples at time of harvest, C and D, show dry weight after tissues were placed in a drying oven at 70˚C until a constant weight was reached. E and F show the % moisture held within the plant; calculated from the fresh and dry weights. A Two-Way ANOVA Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one column, e.g. well-watered/ drought represent significant differences compared to the corresponding WT value.

5.2.8.2.2 Below-ground root biomass

All lines showed significant differences between well-watered and drought-stressed samples (Fig. 5.31B, D & F; Tukey HSD, $P < 0.0001$). When the transgenic lines were compared to wild type in just well-watered or just drought-stressed conditions, none showed a significant difference, but on average *KfCIB1_RNAi_12A* did appear

to have a higher average fresh root weight in well-watered conditions (Well Watered; WT = 8 g, *12A* = 11 g, *30A* = 7 g; Drought; WT = 1.3 g, *12A* = 1.4 g, *30A* = 1.3 g). Once roots had been dried to a constant dry weight, there was also no difference in weight between all well-watered lines, or between drought-stressed lines (Well Watered; WT, *12A* and *30A* all averaged 1 g; Drought; WT, *12A* and *30A* all averaged 0.5g).

All lines showed moisture differences when comparing between well-watered and drought-stressed conditions (Tukey HSD, $P < 0.0001$). However, when comparing between lines, there was no difference moisture content of the roots, suggesting each line maintains approximately the same moisture level in both conditions (Well Watered; WT = 87 %, *12A* = 91 %, *30A* = 85 %; Drought; WT = 63 %, *12A* = 62 %, *30A* = 63 %).

5.2.8.2.3 Induction of UV protective anthocyanin pigmentation in response to drought-stress

Anthocyanin measurements were carried out on WT and both *KfCIB1_RNAi* lines, showing that in drought, significantly more anthocyanins were produced (Fig. 5.32; Tukey HSD, $P < 0.0001$). In well-watered conditions, WT produced a small amount of anthocyanins, but *KfCIB1_RNAi_12A* on average contained more, whilst *KfCIB1_RNAi_30A* produced significantly less leaf anthocyanins (Well Watered; WT = 0.001 μM cm^{-2}, *12A* = 0.002 μM cm^{-2}, *30A* = 0.0002 μM cm^{-2}), and during drought, large increases were seen, although all three lines reached approximately the same concentration (Drought; WT = 0.008002 μM cm^{-2}, *12A* = 0.006002 μM cm^{-2}, *30A* = 0.008002 μM cm^{-2}).

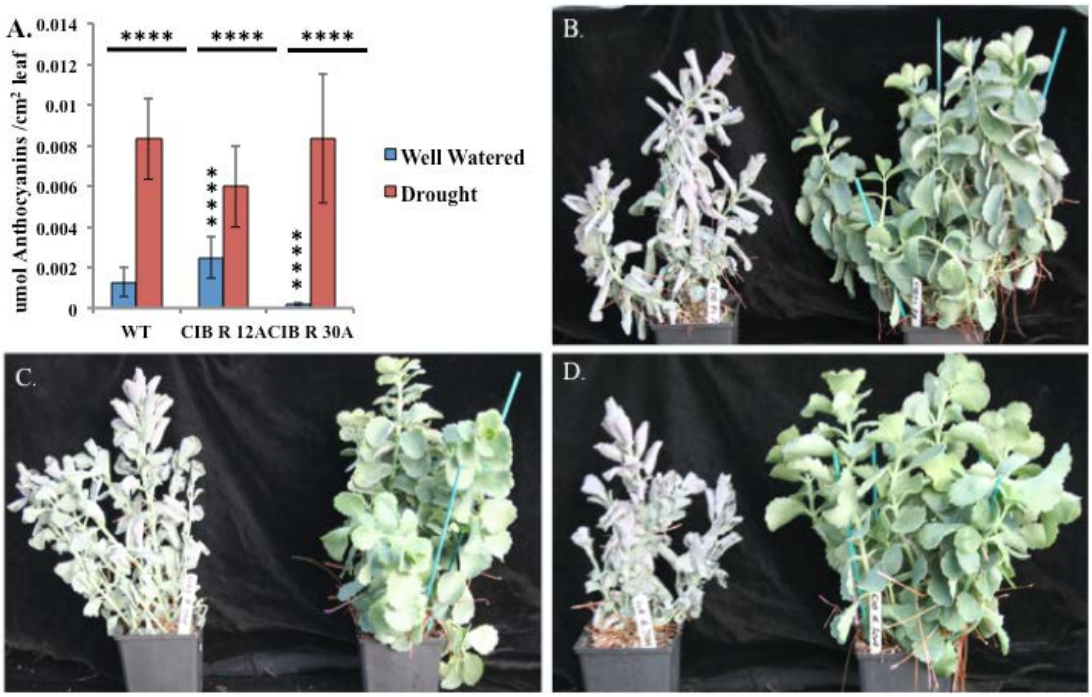

Figure 5.32 . Variation in the level of anthocyanin leaf pigments in response to drought-stress for WT and *KfCIB1_RNAi_12A* and *30A*. **A.** Quantified levels of Anthocyanins in WT and the *KfCIB1_RNAi* lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one column, e.g. drought, represent significant differences to the corresponding WT value. **B.** Images of drought-stressed (Left) and well-watered (Right) WT. **C.** Images of drought-stressed (Left) and well-watered (Right) *KfCIB1_RNAi_12A*. **D.** Images of drought-stressed (Left) and well-watered (Right) *KfCIB1_RNAi_30A*.

In addition, the leaf photosynthetic pigments chlorophyll A and B and carotenoids were also assayed. Chlorophyll A and B, and the carotenoids all increased significantly during drought stress (Fig. 5.33; Two-way ANOVA, $P < 0.0001$). However, when the *KfCIB1_FL* lines are compared to the wild type, they all produced photosynthetic pigments to approximately the same level.

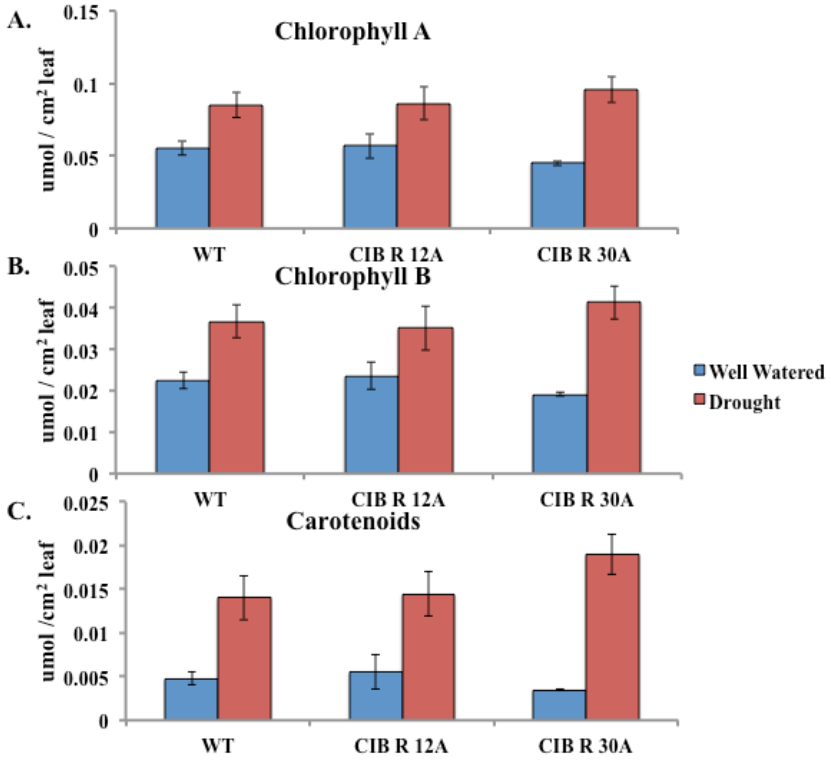

Figure 5.33. Variation in the levels of photosynthetic pigment in response to drought-stress in WT and *KfCIB1_RNAi* lines. **A.** Changes in the levels of Chlorophyll A in response to drought. **B.** Changes in the levels of Chlorophyll B in response to drought. **C.** Changes in the levels of Carotenoids in response to drought. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; **** shows P = ≤ 0.0001. Stars over bars at the top of the graph represent lines that are significantly different between treatments.

5.2.8.2.4 CAM-associated Metabolite levels in drought conditions

In well-watered conditions, malate fluctuations were detected between light and dark for each line ($P < 0.0001$), and this relationship also continued under drought-stressed conditions (Fig. 5.34A & B; $P < 0.0001$). When comparing well watered and drought conditions during the light period, the RNAi lines had significantly less malate at the end of the day (Fig. 5.34C; WT; $P = 0.04$, *12A*; $P = 0.0003$, *30A*; $P = 0.007$). During the dark though, regardless of being in well-watered or drought-stressed,WT still continued to accumulate around 60 µmol of malate per gram fresh weight ($P > 0.999$), whereas both *CIB_RNAi* lines fixed significantly less malate under drought-stress (Fig. 5.34D; 12A & 30A; $P < 0.0001$). This suggests that CAM-idling may have started to occur in the transgenic lines, whereas the wild type was able to continue conducting full CAM after this length of drought.

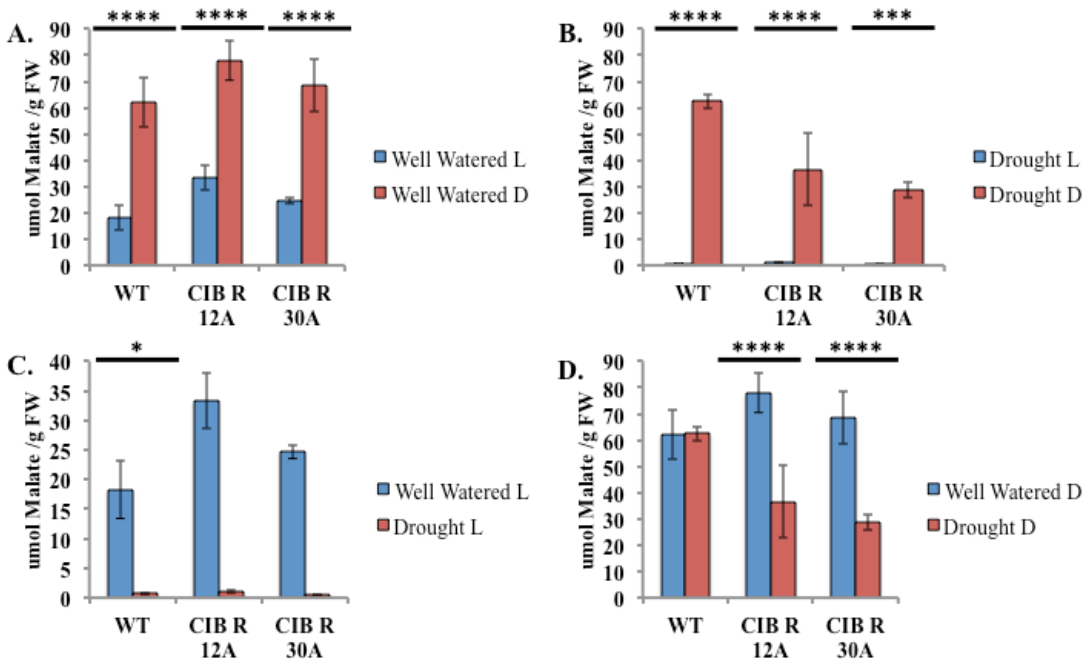

Figure 5.34. Daily fluctuations in malate content of CAM leaves of the WT and RNAi knockdown lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. well-watered and Drought, and stars above one column, e.g. drought, represent significant differences from the corresponding WT value.

5.2.9 *KfCIB1*s impact on reproductive output

K. fedtschenkoi is only able to reproduce by the production of plantlets at leaf margins, and therefore plantlet production from detached mature leaves can be taken as a measure of the reproductive output of the plant. As with all plants, being able to survive and reproduce is the ultimate goal of life, and so it was of interest investigate whether or not the changes in *KfCIB1* gene expression in these transgenic lines had an impact their reproductive output. Both the total number of notches per leaf, and the number of plantlets produced were counted in order to calculate the number of leaf notches to produce one plantlet.

5.2.9.1 The impact of *KfCIB1* over-expression on reproductive output

For the WT leaves, there was an average of 3.1 leaf notches required per plantlet. *KfCIB1_FL_11B* was very similar; with a plantlet being produced every 3.2 notches. *KfCIB1_FL_5B* on the other hand, produced 1 plantlet for every 2.1 leaf notches. This suggested that the high level of over-expression of *KfCIB1* in this line had a positive effect on reproductive output, whilst the lower level of *KfCIB1* over-expression in line *11B* had little effect (Fig. 5.35; One-Way ANOVA, P = 0.05).

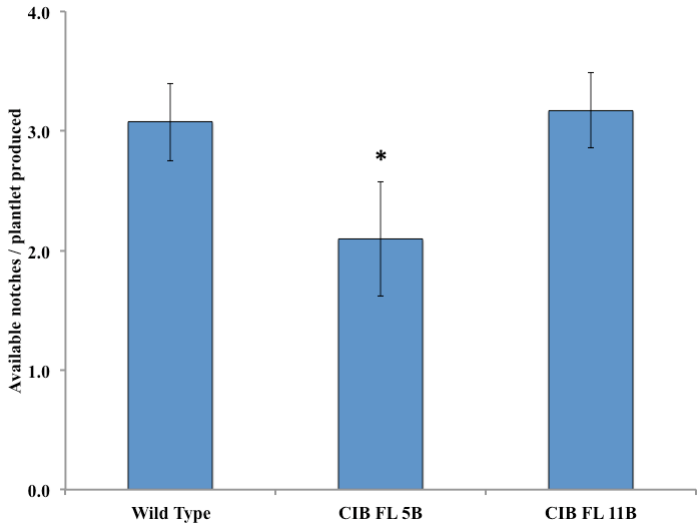

Figure 5.35. Impact of over-expression of *KfCIB1* on reproductive output in the form of leaf margin plantlet formation (For WT n = 39 leaves, 5B n = 20 leaves, for 11B n = 39 leaves). Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars above each column represent significant differences to the corresponding WT value.

5.2.9.2 *KfCIB1* RNAi knockdown lines impact on reproductive success

Both *KfCIB1_RNAi* lines had significantly lower leaf notch:plantlet ratios, than WT who produce a plantlet every 3.1 leaf notches. *KfCIB1_RNAi_12A* produces a plantlet every 2.5 notches (Fig 5.36; One-wayANOVA, P < 0.0001), and *KfCIB1_RNAi_30A* produces a plantlet every 2.7 leaf notches (One-way ANOVA, P = 0.0003). Overall, these results suggest that changes in the transcript level of *KfCIB1* had a positive impact on reproductive output, although it would be important to study the establishment of the plantlets and their ability to produce leaves that were also competent to set plantlets in order to establish whether this increased reproductive output led to an increase in individuals in the next generation.

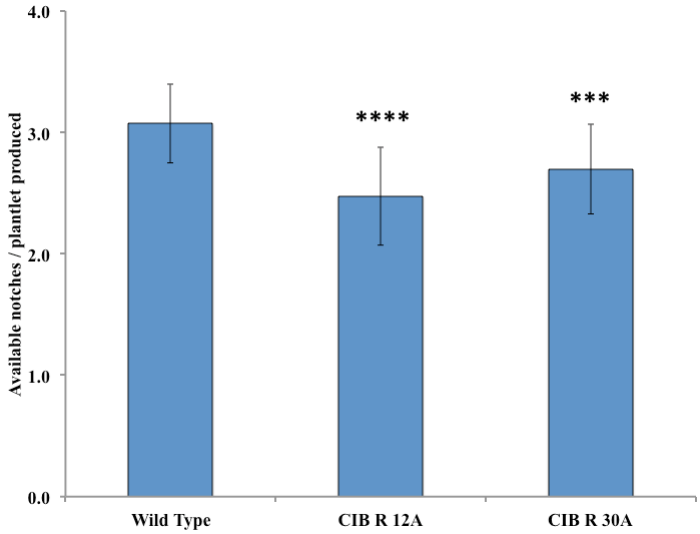

Figure 5.36. Impact of RNAi knockdown of *KfCIB1* on reproductive output in the form of leaf margin plantlet formation (For WT n = 39 leaves, 12A n = 19 leaves, for 30A n = 19 leaves). Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars above each column represent significant differences to the corresponding WT value.

5.3 Chapter 5 – Discussion

This chapter focussed on the investigation of the role of *KfCIB1* in the circadian optimisation of CAM. Initial screening had identified two over-expresser and two RNAi lines and these lines were subjected to detailed phenotypic characterisation in order to learn more about the *in planta* function of *KfCIB1*. Different levels of over-expression and knockdown were chosen, so that any dosage-dependent effects could also be investigated. The four chosen lines: *KfCIB1_FL_5B* and *11B*, and *KfCIB1_RNAi_12A* and *30A*, were subjected to various molecular, biochemical and physiological tests, to try and help define the role of *KfCIB1*.

5.3.1 Bioinformatic analysis of *KfCIB1*

From an amino acid alignment between *KfCIB1* and its closest ortholog in Arabidopsis, *AtbZIP42*, a number of highly conserved residues were identified (Fig. 5.2). Both genes showed identical basic domains, containing the nuclear localisation signal and the DNA-binding sequence, suggesting that they are likely to interact with very similar target DNA-binding sites in promoters of the genes that they regulate. The main difference between the amino acid sequences of the two genes was in the leucine zipper domain, which is responsible for the binding to, and interaction with, the homo- or hetero-dimerization partners of bZIPs (Liu *et al.,* 2014). This therefore suggests these orthologs may bind to similar DNA motifs, but may carry out different functions due to interaction with distinct dimerization partners. This does seem likely, as *bZIP42* is expressed in floral tissues and is involved in petal development and expansion in Arabidopsis, whereas in *K. fedtschenkoi, KfCIB1* in LP6 is clearly carrying out a different function (Riechmann *et al.,* 2000; Jakoby *et al.,* 2002; Ehlert *et al.,* 2006). Again, if *KfCIB1* does play a role in CAM, then it will need to regulate a

whole host of genes that are not under the control of *AtbZIP42* in Arabidopsis, and so changes in the leucine zipper protein-protein interaction domain may have been essential for it to take on these new functions.

When looking in the quantitative RNA-seq data for the other *K. fedtschenkoi* core CAM and clock-gene transgenic lines (*rNAD_ME1*, *rPPDK*, *rPPCK1_3* and *TOC1_OX*), which have had their CAM leaf transcriptome sequenced over the light/ dark cycle, it was clear that the transcript abundance of *KfCIB1* was strongly elevated in *rNAD-ME* and *rPPDK* relative to the wild type, suggesting that *KfCIB1* may play a role in regulating the genes involved in the decarboxylation module of the CAM cycle in the light (Fig. 5.3). Interestingly, levels of *KfCIB1* were greatly reduced in the *TOC1* over-expresser line. This further supports the original proposal that *KfCIB1* is indeed part of the clock output pathway, as it has previously been shown that over-expressing the core clock component *TOC1* led to arrhythmia in output processes and genes (Dall'omo 2011).

5.3.1 The impact of *KfCIB1* mis-regulation of the light/ dark regulation of genes associated with CAM and the clock

For all four *KfCIB1* lines, there were large differences in LD conditions between WT *KfCIB1* transcript levels and the transgenics. *KfCIB1_FL_5B* (Fig. 5.4) showed a 4.3-fold increase, and *11B* showed a 1.4-fold increase (Fig. 5.5). *KfCIB1_RNAi_12A* displayed a large average reduction in transcript level, with 118.7% fold lower than the WT level at peak expression time (22:00; Fig. 5.6), whilst *30A* showed a more moderate reduction of 9.6-fold (Fig. 5.7). Unfortunately, out of all the CAM genes screened with semi-quantitative RT-PCR, there were no differences detected that showed the same magnitude of change as the one measured for *KfCIB1* itself.

However, it is well know that transcript abundance does not necessarily equate to protein abundance, and therefore there may not be the same fold change in in the *KfCIB1* protein as there was in gene transcript level (Maier *et al.,* 2009). Unfortunately, during this PhD, the *KfCIB1* anti-peptide antibody could not be optimised to allow immunoblot detection of the KfCIB1 protein in protein extracts from CAM leaves. From the transcript levels measured in LD cycles, the perturbation of *KfCIB1* transcript levels in these transgenic lines did not affect the core clock genes *KfCCA1* and *KfTOC1*, as neither a reduction nor increase in transcript level was detected, suggesting that *KfCIB1* did function more in output pathways from the clock.

More subtle differences were detected for a number of genes. For example, *KfPPCK1* displayed a slight increase in transcript level at peak expression compared to wild type for both *RNAi* lines, with the largest increase being seen in 12A, the strongest knockdown line (Figs. 5.6 & 5.7). In the FL lines, both show a decrease in *KfPPCK1* expression in the dark, with the strongest overexpresser showing the largest decrease (Figs. 5.4 & 5.5). This is interesting as it had been demonstrated previously using electrophoretic mobility gel-shift analysis that *KfCIB1* could bind to putative ACGT-motifs in the *KfPPCK1* promoter (Meszter *et al.,* 2008; Meszter 2010). However, if *KfCIB1* does bind to and regulate the *KfPPCK1* promoter then it is surprising that when *KfCIB1* was knocked down, *KfPPCK* levels were not decreased to a similar magnitude. Overall, these gene transcript level results do support the proposal that changes in *KfCIB1* levels impact either directly or indirectly on *KfPPCK1* levels.

All lines also displayed differences in *KfNAD-ME α* and *KfNAD-ME β* transcript levels, with the *KfCIB1* over-expressers showing the largest changes (Figs. 5.4 & 5.5). In *KfCIB1_FL_5B,* there was an enhanced oscillation of *KfNAD-ME α* and *KfNAD-*

ME β transcript levels, whilst in *KfCIB1_FL_11B*, there were constantly increased levels. There were smaller changes for the RNAi lines. There were also changes in the transcript levels of genes associated with starch breakdown and the movement of sugars and sugar-phosphates between the cytosol and the chloroplast. These included *KfGPT2, KfMEX1* and *KfGWD*. In addition, *KfPPDK-RP* also displayed small changes in its transcript abundance. This again provides more evidence that *KfCIB1* may play a role in the circadian optimisation of CAM rather than it targeting only one downstream gene, as it had a widespread, if often rather subtle and limited, effects on potential downstream target genes.

Two *KfCIB1* transgenic lines were also assayed under LL free-running constant conditions in order to determine whether there was a perturbation of circadian rhythms of clock and CAM gene transcripts due to the over-expression or reduction of the *KfCIB1* gene. *KfCIB1_FL_11B* and *KfCIB1_RNAi_30A* were used for this experiment. Plants were sampled whilst in LD conditions at the end of light and dark (10:00 and 22:00 ZT), to show correct gene cycling in these conditions. Once transferred to LL they were left for a day to enter true free running conditions, and then at 26 h LL, the first LP6 sample was collected. The last sample was collected at 90 h LL. The period length of the free-running rhythm of CO_2 fixation, and the oscillations of CAM and clock gene transcripts in *K. fedtschenkoi* under the LL conditions used in this experiment is approximately 18 - 21 h (Dever *et al*., 2015), so it was expected that in LL, the genes may not peak at their predicted times in 'subjective light' or 'dark'(Luttge *et al.,* 1992).

Interestingly though, under LL conditions, it took over 72 h before *KfCIB1* started to peak out of time relative to the wild type in the transgenic lines. In *KfCIB1_FL_11B*, as well as maintaining a *KfCIB1* rhythm at subjective dark, this line also had a knock-on effect that caused both *KfCCA1* and *KfTOC1* to continue peaking in phase with their timing under LD cycles, despite the endogenous rhythm being less than 24h (Fig. 5.8).

With *KfCIB1_RNAi_30A,* the *KfCIB1* transcript level was very interesting. Despite these plants carrying in their genome a hairpin RNAi construct, which caused a dramatic and near complete reduction in transcript levels in LD cycles (Figs. 5.7 & 5.9), once the line was transferred to constant conditions, this was somehow over-ridden, and levels of *KfCIB1* transcripts were actually higher than wild type (Fig. 5.9). This could therefore mean that *KfCIB1* is important in the maintenance of circadian rhythms through environmental change, such as increased light, which is one of the main functions of group S of bZIP TFs (Jakoby *et al.,* 2002). Despite this weak over-expression of *KfCIB1* in LL in the *RNAi_30A* line, the impact on the clock genes *KfCCA1* and *KfTOC1* was not the same as observed with the *FL_11B* over-expresser (compare Fig. 5.8 with Fig. 5.9). *KfCCA1* continued to peak at the end of 'subjective dark' for 50 h under LL, which was later than the peak in the WT. After 50 h, the period began to shorten. Also, after 50 h under LL conditions, *KfTOC1* transcript oscillations collapsed towards arrhythmia in *RNAi_30A*, whilst the wild type continued to cycle (Fig. 5.9). Neither the *FL_11B* nor the RNAi line showed much difference compared to the wild type in terms of the LL transcript abundance oscillations of *KfPPCK1* expression though, except that there was a more rapid fall in the amplitude of the oscillations, and in the last 24 h the *KfPPCK1* transcript

collapsed towards arrhythmia, most noticeably in *FL_11B*. This does suggests that although *KfCIB1* has been shown *in vitro* to be able to bind to ACGT-motifs found in the *KfPPCK1* promoter, in reality it may not bind to a regulate these motifs in plant, or at the very least it may be redundant with other TFs and/ or may need to interact with other TFs in order to have its full impact on the regulation of downstream targets.

This is why HA-Tagged *KfCIB1_FL* lines were produced in *K. laxiflora* in the final year of this project. Although there was not enough time to undertake ChIP-Seq using these lines within this project, they will in future allow the discovery of exactly where this transcription factor binds to the *K. laxiflora* genome. Alternatively, the specific heterodimer needed to bind and elicit the correct control of *KfPPCK1*, may not have formed in the over-expresser or RNAi lines of *KfCIB1* studied here; possibly due to the level of the other bZIP heterodimerisation partner not being sufficient to form the active dimer in the *FL* lines, or the dimerization partner may be able to homodimerise in the absence of *KfCIB1* in the RNAi lines, thereby rescuing most of the functions regulated by KfCIB1 itself. One potential dimerization partner of *KfCIB1* could be *KfCIB2*, which was a second bZIP factor identified previously in the Hartwell lab by a former PhD student, Roland Meszter. This bZIP possessed the same basic domain as *KfCIB1*, and an eight heptad repeat leucine zipper domain, also placing it in group S (Meszter 2010). It has also been suggested that bZIPs tend to dimerise with functionally related bZIPs (Weltmeier *et al.,* 2006; Correa *et al.,* 2008), and so *KfCIB2* seems like another interesting candidate to investigate for the circadian control of CAM.

5.3.2 Gas exchange rhythms

The alterations to *KfCIB1* transcript levels and its light/ dark regulation in these transgenic lines did have a large impact on the physiological gas exchange rhythms of the transgenic lines. Both the over-expresser and RNAi knockdown lines possessed extremely sensitive stomata, which adapted quickly to changes in the light conditions, for both single detached leaves with their petioles in distilled water, and a leaf attached to a well-watered plant.

At the end of the dark, *KfCIB1* lines behaved in a similar way to the wild type, as stomata closed due to the commencement of malate decarboxylation. Malate decarboxylation leads to CO_2 being released inside the leaf, producing a high partial pressure (pCO_2), which helps with water conservation (Maxwell *et al.,* 1999).

Once there is a drop in this pCO_2 under well-watered conditions, then the stomata re-open. It appears that when pCO_2 drops to a low enough level, the wild type opened its stomata slightly, which helps with water conservation, whereas the *KfCIB1* lines appear to have lost this fine control, and open their stomata much wider (Figs. 5.10 – 5.14). This could be one of the reasons why when these lines were subjected to drought-stress, the majority of the *KfCIB1* lines (except *FL_11B*) appeared to go into CAM-idling, perhaps because they lost water more rapidly through transpiration due to this excessive opening of stomata**.**

It also appears these *KfCIB1* transgenic lines get the 'best of both worlds' where their stomata appear to be open during the light and dark phases, and so this could explain the increased levels of total sugars and the trend towards larger daily variations in starch levels, but as growth did not improve in these lines, it seems likely that futile cycling of metabolites may be occurring. It has been shown that a drop in pCO_2 is a signal for stomatal opening (Bohn *et al.,* 2001), but in the *KfCIB1* plants, lots of CO_2

is released at dawn, suggesting that in these lines, pCO_2 may not be sensed correctly. It could also suggest that in CAM plants, this pCO_2 sensing is gated, where during the day it can lead to re-opening of stomata for phase IV if the plant is well-watered, but during the dark period, stomata will open regardless of water status in order to ensure CAM is carried out efficiently.

As there are no significant differences found in size or growth between the transgenic lines and wild type, then this could indicate the leaf carbohydrates are simply being turned over due to futile cycling. More detailed time courses and investigations into levels of other primary metabolites could be very informative to determine what the elevated pools of soluble sugars are used for.

5.3.3 Metabolite levels

Malate, starch and sugar levels were quantified in these lines to see if *KfCIB1* over-expression or knockdown impacted on metabolite levels. Malate and starch cycled in a reciprocal fashion between the end of the light and the end of the dark. Starch can be broken down to produce sugars, which can then be used for various processes, such as plant growth and/ or the production of carbon skeletons for PEP synthesis during Phase I (Haider *et al.,* 2012).

Malate and starch levels underwent diurnal changes between dawn and dusk in each line; both *KfCIB1_FLs* and *KfCIB1_RNAis*. Malate levels for the WT and all four transgenics were approximately equal at dawn, suggesting malate accumulation was not affected by the perturbations of *KfCIB1* (Figs. 5.15 & 5.19). For starch levels though, *KfCIB1* levels did have an impact. When compared to the wild type, the *FL* lines possessed on average more starch during the light period, but at the end of the

dark, there were no differences compared to the wild type (Fig. 5.16). On the other hand, the *KfCIB1_RNAi* lines showed small reduction in starch levels at the end of the light and dark periods (Fig. 5.20). It has been demonstrated that a reduction in carbohydrates leads to a reduction in CO_2 uptake, which can then have a knock on effect to increase *PPCK1* expression, which we do see to small extents in transcript levels at the end of the dark (Dodd *et al.,* 2003).

Both *KfCIB1_FL* lines and *RNAi_30A* metabolised more of their stored starch during the dark period than the wild type. With the leaf soluble sugars, there were on average significantly more total sugars (glucose, sucrose and fructose) in all of the transgenic *KfCIB1* lines (Figs. 5.18 & 5.22), relative to the wild type, which suggests either that more starch was being broken down, or that there was a problem with starch synthesis or futile cycling of metabolites is occurring due to changes in the circadian control of CAM. It has been proposed in the literature for a while now that the metabolic products of photosynthesis, including sugars, are able to act as checkpoints for the integration of signals from the circadian clock, in order for it to modulate CAM in response to the ever-changing environment, thereby facilitating CAM's well-known photosynthetic plasticity (Borland *et al.,* 2004).

These excess sugars in the plant can also be used for a variety of processes, such as growth, increased PEP production, and in *Clusia* species can feedback and affect *PPCK* diel cycling (Borland *et al*., 2004; Borland & Taybi, 2004). Furthermore, sugars have also been shown to be accumulated by plants that are under stress conditions, which could indicate that changes in *KfCIB1* expression levels are causing a stress response in the plant (Corrales *et al.,* 2014). Sugars, especially glucose, have also been shown to down-regulate genes involved in biotic and abiotic stress,

carbohydrate and protein synthesis, which adds another level of complexity to the challenge of determining the effect of this TF, rather than that of sugars.

5.3.4 Protein levels and enzyme activity

Firstly, PEPC and phospho-PEPC protein levels were determined. All lines displayed no differences in PEPC levels, which also indicated equal loading of the SDS-PAGE gels used for the immunoblots. For the wild type, phosphorylation of PEPC was detected throughout the dark period, with the highest level of phosphorylation detected at 6 h and 10 h into the 12 h dark period. Both *KfCIB1_FL_5B* and *KfCIB1_RNAi_30A* showed differences in PEPC phosphorylation state. *5B* showed no phosphorylation at the start of the dark period (Fig. 5.23B), and *30A* showed a clear decrease in the phosphorylation state of PEPC at 10 h dark, 2 h before dawn (Fig. 5.24B). This therefore suggested that both of these lines would contain PEPC that was more sensitive to malate inhibition at these time points, which would likely result in less CO_2 being taken up and fixed as malate (Borland *et al.,* 2004).

When PEPC activity was assayed for total extractable activity, the results revealed that all four *KfCIB1* lines had increased activity compared to the WT when expressed per mg protein, which may be a way of compensating for lower phosphorylation (Fig. 5.25A). Furthermore, it may be due to more PEP being present in the enzyme samples, if that is what the excess sugars are being used for, from increased nocturnal starch breakdown. When PEPC activity was expressed per mg chlorophyll though, all lines showed a slight decrease in PEPC activity level relative to the WT (Fig. 5.25B). This could therefore indicate total protein changes in the leaves of these transgenic lines, whilst chlorophyll levels have been shown to change very little in these lines (Figs. 5.29 & 5.33).

As phosphorylation can make PEPC up to ten times more resistant to malate inhibition, PEPC malate sensitivity assays were conducted. These showed that all of the transgenic lines contained PEPC that was more sensitive to malate inhibition that the enzyme in the wild type leaves at the same time points. This suggested that all 4 lines had differences in PEPC phosphorylation compared to WT, as supported by the immunoblots for *FL_5B* and *RNAi_30A* (Tables 5.5 & 5.6).

Levels of other key CAM proteins were also measured using immunoblotting. NAD-ME α and NAD-ME β levels were slightly higher at the start of the light and dark periods in *KfCIB1_FL_5B* (Fig. 5.23B & C), which fits with the transcript abundance data for the corresponding genes (Fig. 5.2). *KfCIB1_RNAi_30A* on the other hand, had lower levels throughout the dark period (Fig. 5.23B & C).

5.3.5 Performance in drought

As CAM is an adaptation to drought, it was important to identify whether changes to expression of this transcription factor led to changes in the ability of these plants to survive in harsh environments. Plants were grown for 6 months in greenhouse conditions, and were then split into different growth conditions, either watered daily or not watered at all for 31 days.

Overall each line showed large significant differences when comparing drought and well-watered conditions for all measurements; fresh weight, dry weight and moisture content. When comparing these factors between lines in just the drought or well-watered conditions, then there were no large differences. The most interesting results involved the dry weights of above ground material for *KfCIB1_FL_11B, RNAi_12A*

and *RNAi_30A*, which showed no statistical difference between well-watered and drought conditions, suggesting that these lines were able to grow and fix as much carbon under either condition. However, the amount of moisture contained in the plant did differ, leading to significantly different fresh weight measurements (Figs. 5.27 & 5.31).

During drought conditions, plants are under a great deal of stress. In *K. fedtschenkoi,* a key visible response to severe drought stress is the development of a purple hue by the leaves, which is due to the production of anthocyanins. The photosynthetic pigments chlorophyll A and B, and carotenoid were also measured. All of these pigments were seen to increase in concentration under drought-stressed conditions, even after leaf area variations between well watered and drought were taken into consideration. The *KfCIB1_FL* lines had consistently lower levels of anthocyanins in both conditions (Fig. 5.28), whilst the RNAi lines were very similar to the wild type (Fig. 5.32). For chlorophyll and carotenoid levels, although they did all increase with drought, there were no large differences when comparing the transgenic lines to the wild type (Figs. 5.29 & 5.33).

To ensure that CAM was still occurring malate levels were quantified to check whether it continued to cycle between light and dark at the end of the drought-stress treatment. The majority of the lines had begun to carry out CAM-idling, keeping their stomata shut both day and night, and re-fixing respired CO_2 behind closed stomata. CAM-cycling is accompanied by the continued cycling of malate levels between dawn and dusk, but at much lower levels than when direct atmospheric CO_2 fixation is occurring in the dark period. Only the wild type and *KfCIB1_FL_11B* were likely to

be still conducting full CAM, because both lines showed significant differences in malate concentration between dawn and dusk when drought-stressed. Furthermore, both of these lines also showed no significant difference between well-watered and drought-stressed conditions in terms of their malate levels at the end of the dark period (Fig. 5.30D). Overall, the results obtained from the drought-stress experiment did support the conclusion that mis-regulation of the *KfCIB1* transcript level resulted in a reduced ability of the plants to carry out efficient and optimised CAM when in a water limited environment. In particular, these lines were more prone to switch to CAM-idling earlier on during drought-stress. This may well be an advantage, as the physiological significance of CAM is to conserve water and carbon in response to environmental restrictions, and these *KfCIB1* lines were able to respond to soil water drying more quickly than the wild type, which may mean that they would be able to survive drought for longer than the wild type but not grow as productively in short term drought (Borland *et al.,* 2004).

An interesting follow-on experiment would be to identify how quickly these lines entered into CAM-idling, and if that had an effect on growth and water content compared to wild type. This could be studied in detail using a gas exchange system such as the LICOR 6400XT which is able to make accurate measurement of the stomatal conductance.

5.3.6 Reproductive output

A key hypothesis as to why CAM has evolved in many diverse plant species that are adapted to arid and semi-arid environment states that the evolution of CAM in these lineages gave them a selective advantage in terms of competition and survival in the seasonally dry environments in which they are mostly found. A key step in the

selection of any advantageous trait during evolution is that it is passed to the next generation, and the next generation compete and reproduce better than their progenitors.

K. fedtschenkoi has lost the ability to produce viable dry seed because the developing seed fail to achieve dehydration tolerance in the final stages of seed set. The same mutation that affects seed dehydration tolerance is also believed to underpin the ability of *K. fedtschenkoi* leaves to produce adventitious plantlets from the notches along the margins of its leaves. Three out of the four KfCIB1 plants, namely *FL_5B,* and the *RNAis_12A* and *30A*, showed an increase in reproductive output compared to the wild type, producing more adventitious plantlets per leaf margin notch on average (Figs. 5.35 & 5.36). Interestingly, the line that performed as well as the wild type; *KfCIB1_FL_11B*, was also the line that seems most 'wild type-like' in several other experiments, such as in the drought experiment (Fig. 5.35). This could therefore indicate that if CAM is impaired in some way in *K. fedtschenkoi*, this can lead to the plants investing less energy into the maintenance of CAM and more in producing the next generation. Furthermore, in order to calculate reproductive output, leaves were detached from mature adult plants (> 6 months old), which would mean that each leaf was undergoing drought stress when producing plantlets from the notches, due to detachment.

5.4 Conclusions

From in-depth analysis of the phenotypes of the *KfCIB1* transgenic lines in which the transcript level of the genes was either up- or down-regulated, it is clear that this CAM-induced bZIP transcription factor may well play a role in the circadian optimisation of CAM. Evidence suggests *KfCIB1* may control genes involved with stomatal control or development due to the rapid fluctuations in stomatal aperture seen using the LICOR as light turn off or turn on, but also is likely to indirectly control many CAM genes by being part of the clock output pathway. This conclusion was supported by the LD time course data, where there were widespread but small knock-on effects on the transcript abundance of a range of CAM and clock genes. Furthermore, under LL conditions, transcript level oscillations of both the clock genes *KfCCA1* and *KfTOC1* and the clock-controlled CAM gene *KfPPCK1* showed pronounced circadian phenotypes. One particularly bizarre and unexpected result was that the RNAi line was unable to maintain silencing of the *KfCIB1* endogenous transcript level during in free-running LL conditions. This could suggest that the 35S promoter is being downregulated in LL, or possibly the DICER RNAi pathway may be somehow switched off in constant conditions.

Gas exchange rhythms, using both the multi-channel system for detached LP6, and the LICOR from a single LP6 leaf attached to a well-watered plant, both demonstrated large spikes of CO_2 fixation and/or release the light to dark transition. The LICOR enabled the accurate measurement of stomatal conductance, showing that as well as these CO_2 spikes, there were also rapid changes in stomatal aperture. Changes to *KfCIB1* expression also led to significantly increased soluble sugar levels, which could suggest these transgenic plants are under stress conditions. Malate accumulation in the dark period appeared to be slightly increased in these lines,

although they were not significantly different to the wild type. This seems odd, as the K_i of PEPC for malate in all of these lines was found to be reduced compared to the wild type, and the phosphorylation state of PEPC was also reduced, and yet, both RNAi lines show a small increase in dark period CO_2 fixation of just over 3% (Table 5.4). Starch levels were also affected by perturbations of *KfCIB1* levels, as the over-expresser lines showed an increase in starch accumulation during the light, whereas the *KfCIB1_RNAi* lines showed a decrease in starch both at the end of the light and dark periods.

Under drought-stress conditions, when the plants ability to conduct CAM can become more critical to its survival, both the *FL* and RNAi lines showed small increases in dry weight matter. This therefore suggests these plants were able to achieve a higher WUE and also were able to grow more productively in water-limited conditions in short-term drought. This may be because these lines entered CAM-idling earlier than the wild type as soil drying increased following the with-holding of watering. If these lines are no longer opening their stomata then they are able to conserve more water. This increase in dry weight was only detected in shoots though, not in below ground material sampled.

KfCIB1 did not really impact on chlorophyll or carotenoid levels, but did have an on the level of anthocyanins in the leaves. The over-expresser lines showed a decrease in anthocyanins under both well-watered and drought-stressed conditions, and the RNAi lines also showed differences. This may be how plants are able to grow more productively in drought, if more energy is directed to growth rather than stress responses.

Changes in the expression of *KfCIB1* also led to a higher proportion of leaf notches producing plantlets along the leaf margins in these transgenic lines, which suggests

that adjusting the expression of this TF caused the diversion of resources towards plantlet production, perhaps at the expense of retaining the ability to perform full CAM.

When attempting to define the function of *KfCIB1* in CAM, there are some important points to consider. When trying to determine the function of any bZIP transcription factor, is that it will only bind to DNA if it has formed a dimer. The effects of the mis-expression of *KfCIB1* on CAM discovered in this work are most likely due to changes in the level of the homo-dimer, especially in the over-expresser lines where any heterodimerisation partner of *KfCIB1* would not be present in sufficient abundance to allow for the formation of an excess of active heterodimers. Thus, although the transcript level for *KfCIB1* was increased in the *FL* lines, it remains possible that it must form a heterodimer with a different bZIP in order for it to be able to control CAM in a way that would have a larger impact on the physiology and metabolism than the results presented here displayed. It will therefore be important in future work to mine the existing genome and transcriptome datasets for *K. fedtschenkoi* and *K. laxiflora* in greater detail in order to identify other potential bZIPs, which may heterodimerise with KfCIB1. One clear candidate for this function is *KfCIB2* (KF23055), which was found previously to show weak CAM-induction and oscillate in abundance over the light/ dark cycle (Meszter, 2010). The draft genome sequence of *K. fedtschenkoi* also contains a third, completely uncharacterised group S bZIP (KF29910) that is also closely related to *KfCIB1*, so it will be important to generate double and triple over-expression lines in the future that co-express *KfCIB1* and *KfCIB2*, *KfCIB1* and *KfCIB3*, and *KfCIB1, KfCIB2* and *KfCIB3,* as well as double and triple RNAi lines, or perhaps CRISPR/ Cas9 guided double and triple mutants.

Finally, in order to develop a more detailed view of the downstream genes that *KfCIB1* binds to and regulates, HA-tagged overexpressing lines were generated in *K. laxiflora* in the final year of this project. These lines will hopefully allow ChIP-Seq experiments to be conducted, enabling the identification the downstream promoters that KfCIB1 binds to in the genome. Such as genome-wide picture of which genes have KfCIB1 bound to their promoters at different times of day and at different developmental stages of the leaves will allow the identification of all of the biological processes, including facets of CAM, that are regulated by *KfCIB1*.

Chapter 6 – Characterisation of the CAM-associated phenotypic changes in transgenic RNAi and over-expression lines of *K. fedtschenkoi* that are perturbed for the CAM-induced and clock-controlled transcription factor *CYCLING DOF FACTOR2*

6.1 Introduction

As described in Chapter 3, the transcription factor *CYCLING DOF FACTOR2* (*KfCDF2*, KF17950 in the *K. fedtschenkoi* draft genome) was identified in initial screens of RNA-seq data for CAM-induced and light/ dark regulated transcription factors. *KfCDF2* is the closest ortholog in the *K. fedtschenkoi* genome for the Arabidopsis gene *CDF2* (At5g39660). This gene belongs to a family of DNA-binding with one finger (*DOF*) transcription factors that are found exclusively in plants (Riechmann *et al.,* 2000), and they have been shown to play roles in a wide variety of processes. Some of the known roles include phytochrome-regulated expression (Park *et al.,* 2003), guard cell specific expression (Plesch *et al.,* 2001; Negi *et al.,* 2013) and flowering time regulation (Fornara *et al.,* 2009; Corrales *et al.,* 2014). Therefore the *DOF* family of transcription factors was an interesting family to investigate further in the context of the circadian control of CAM. Specifically, published findings from Arabidopsis demonstrate that members of the *DOF* family are important for the circadian control of gene expression, and also accurate temporal stomatal control, both of which are critical points for optimised control within the daily cycle of CAM.

6.1.1 *DOF* family transcription factors

Many *DOF* transcription factors carry out completely unrelated functions, with the only similarity being that they all are linked to moderating gene expression in plants. Arabidopsis is predicated to have 36 DOF TF genes, whilst most other species have been shown to have lower numbers, such as barley with 26 and rice with 30

Chapter 7 – An investigation of short-day dependent induction of flowering in *K. fedtschenkoi* and its perturbation by transgenic manipulation of the *CYCLING DOF-FACTOR2* transcription factor

7.1 Introduction

The *KfCDF2* over-expressing lines described in Chapters 3 and 6, were found to transition to flowering under long-day 16: 8 light: dark greenhouse conditions, whereas the wild type has never been observed to flower under long-days. It was reported in Chapter 3, that the more strongly over-expressing *KfCDF2* transgenic lines produced flowers in long days, whilst several lines that had lower levels of over-expression only produced hooked meristems; an early sign of the transition of the shoot apical meristem towards becoming a floral meristem. This therefore suggested that *KfCDF2* plays a role in the regulation of flowering time, which is a process that is known to be tightly regulated by the circadian clock. As previously mentioned, the most thoroughly characterised ortholog of *KfCDF2* is called *CDF2* in Arabidopsis, for which many papers have been published. *DOF* TFs, namely *CDF1*, *2*, *3* and *5*, have all been shown to function somewhat redundantly in flowering induction in Arabidopsis (Fornara *et al.,* 2009). The phenotype after genetic manipulation in *K. fedtschenkoi* and Arabidopsis was found to differ, likely because Arabidopsis is a long day flowering plant (LDP), whilst *K. fedtschenkoi* is a short day plant (SDP).

The Arabidopsis *cdf2* mutant lines displayed early flowering compared to wild type in both long and short days (Fornara *et al.,* 2009), whilst constitutive overexpression using the CaMV35S promoter led to a delay in flowering, but only in long days (Yanagisawa 2002). From the initial screening results presented in Chapter 3, *KfCDF2* overexpression was only found to cause early flowering in long days, when these plants normally do not flower, so it was of interest to identify phenotypes when the plants were grown under different day lengths, to investigate whether the *K.*

fedtschenkoi CDF2 gene performs the reverse function relative to the orthologous gene in Arabidopsis.

7.1.1. Flowering time regulation in higher plants

7.1.1.1 Regulation of flowering time in Arabidopsis via the day-length dependent pathway (long day flowering)

Flowering in Arabidopsis can be initiated by four genetic pathways: gibberellin (GA), autonomous, vernalisation and light-dependent (Komeda 2004). The strict regulation of the timing of flowering is believed to be under this range of regulatory pathways because reproduction requires a lot of plant resources, and so plants need to ensure that they flower when the conditions are optimal, both in terms of their own development and the external environment and so flowering has been demonstrated to be a repressible process (Komeda 2004; Suzuki *et al.,* 2006). It has long been established that day length is perceived by mature leaves (Knott 1934), which results in a mobile flower-inducing signal called 'florigen' being produced (Chailakhyan 1936). The florigen then moves from the leaf to the meristem (Zeevaart 1976). In Arabidopsis, the gene *FLOWERING LOCUS T* (*FT*) has since been identified as possessing the properties of a mobile florigen flower-inducing signal (Huang *et al.,* 2005). For FT to be able to transmit this flowering signal the protein must move from leaf companion cells into the phloem and then to the shoot apex (Corbesier *et al.,* 2007; Mathieu *et al.,* 2007). The ER-localised protein FT-INTERACTION PROTEIN1 (FTIP1) enables export of FT from the companion cells to the phloem (Liu *et al.,* 2012). FT is then able to travel to the meristem, where FD is being expressed and the FT-FD complex is then able to induce other genes essential for

flowering (Abe *et al.,* 2005). Therefore after *FT* expression, the gene changes related to flowering are most likely located in and around the meristem.

Monitoring seasonal changes in both the quality and quantity of light is a key method that plant use to determine the time of year. The changing properties of a plant's light environment are interpreted via the circadian clock, and signals generated as a result of changing light quality and quantity can induce flowering via an output pathway from the central clock. *GIGANTEA* (*GI*) is a key component of the clock which also plays a critical role in linking the core clock to the light-dependent flowering signalling cascade that can lead to flowering (Fig. 7.1). GI induces the expression of *CONSTANS* (*CO*), a Zinc finger domain protein. *CO* is also up-regulated by the autonomous pathway, temperature, GA and vernalisation cues (Pose *et al.,* 2012). The strength of this signal is proportional to the earliness of flowering. *CO* is regulated by other Zinc finger domain proteins: *CYCLING DOF FACTORs 1, 2, 3* and *5* (*CDF1,2,3&5*). These are able to repress *CO* and prevent flowering.

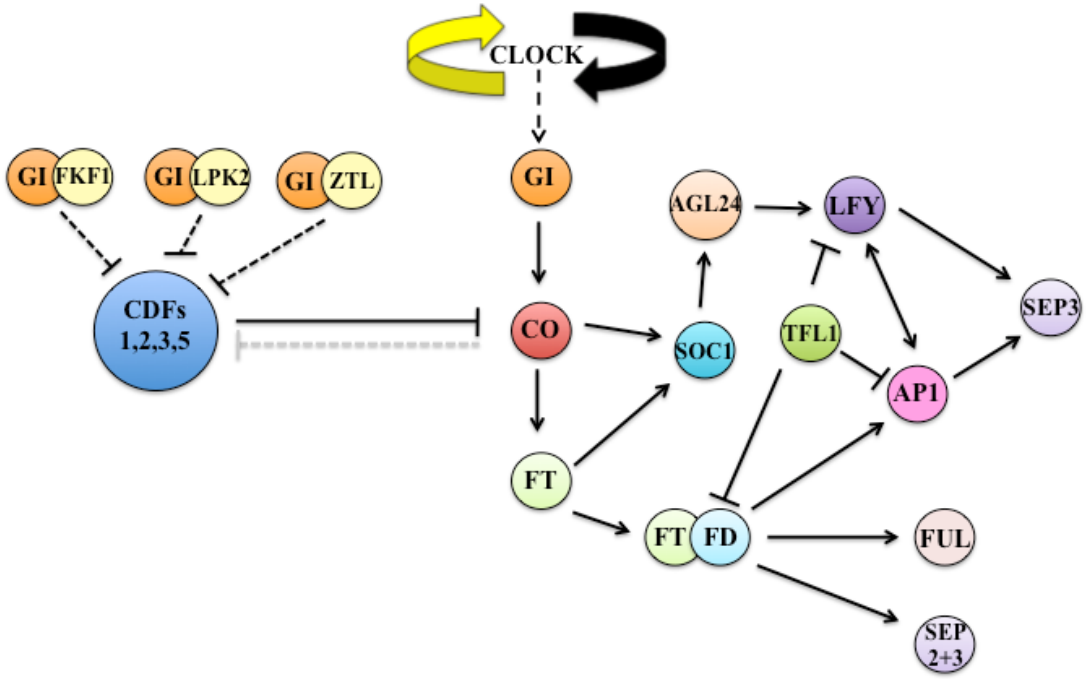

Figure 7.1. Schematic diagram showing an outline of the basic photoperiod-dependent flowering time pathway in the long-day plant *Arabidopsis thaliana*. GI sends the signal from the clock when the light period is long enough to induce flowering. This in turn induces *CO*. CO induces *SOC1* and also *FT*, which then go on to induce many other flowering genes. Once *AP1*, *SEP3* and *FUL* are expressed in the meristem, flowering is initiated. Following the activation of these genes, different classes of genes are activated in different whorls of the developing floral meristem to define and control floral organ development.

In Arabidopsis, once *CO* expression levels reach a threshold that overlaps with the end of the light period in long-day conditions, it is thought to interact with other proteins in order to bind to the promoter of and activate transcription of *FLOWERING LOCUS T* (*FT*) and *SUPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*). FT is then able to move through the phloem from the leaf to the meristem, where FT protein interacts with FD. This protein complex can then further activates the expression of the TF SOC1, which then forms a complex with AGAMOUS-LIKE24 (AGL24), which can then promote the expression of *LEAFY* (*LFY*) and *APETALA1* (*AP1*), which are floral meristem identity genes (Yoo *et al.,* 2005). Increased expression of *SOC1* and *AGL24* is often referred to as the key step regulating the

transition from vegetative to inflorescence meristem development (Wellmer *et al.,* 2010).

In later stages of the floral induction process, *SOC1* and *AGL24* are repressed, to prevent reversion of flowering (Lee *et al.,* 2010). As part of the flowering checkpoints, which fine-tune the balance-point between the development of the inflorescence meristem and floral meristem, the FT homolog TERMINAL FLOWER1 (TFL1) acts antagonistically to FT, and so can repress *LFY* and *AP1* expression, and therefore prevent flowering progression. However, once sufficient AP1 has been produced, which produces a strong flowering signal, it can in turn repress *TFL1* and enable flowering to progress efficiently.

Once FUL, SEP3 and AP1 accumulate in the meristem, the inflorescence to floral meristem transition has begun (Teper-Bamnolker *et al.,* 2005).

This high expression of *AP1* also begins the floral meristem development, where sepals, petals, stamen and carpels are formed, by via the ABCDE model of floral organ identity development (O'maoileidigh *et al.,* 2014).

7.1.1.2 Floral meristem identity genes

Proper development of floral organs at the apical meristem is maintained by five classes of floral organ identity genes found in different amounts within the four whorls: A, B, C, D and E (Fig. 7.2). The proteins in class A are AP1 and AP2. They are able to repress C class genes, enabling sepals and petals to form. B class genes are *AP3* and *PISTILLA* (*PI*), which are involved in the formation of petals and stamens, and C class gene *AGAMOUS* (*AG*) is involved in carpel formation. Group D genes: *SEEDSTICK* (*STK*) and *SHATTERPROOF1* and *2* (*SHP1* & *2*) are responsible for the development of ovules within the carpels (Galimba *et al.,* 2015), and group E genes,

SEP1,2,3 and *4,* function as co-regulators in all whorls. The up-regulation of *AP1* signals the start of floral meristem determination. The floral meristem is also maintained by the positive feedback loop between AP1 and LFY (Pose *et al.,* 2012; O'maoileidigh *et al.,* 2014).

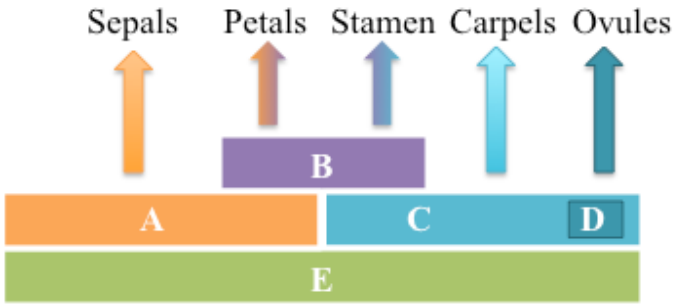

Figure 7.2. Schematic diagram showing the four gene classes involved in floral organ determination. Different interactions between the classes leads to different floral organs being produced by different parts of the floral meristem.

At the beginning of flower development, defined as stages one and two at the meristem, the primordium proliferates without differentiation, which is achieved by the silencing of *SEP3* by SOC1, ALG24 and SHORT VEGETATIVE PHASE (SVP), which interacts with TFL1. At stage three, differentiation begins, with formation of the outer whorl: sepals. At this point *SEP3* expression is promoted by AP1 and LFY, which can then go on and activate B and C class genes in the three inner whorls, which in turn leads to *SOC1* repression to enable the plant to fully commit to flowering.

SEP3 induction also leads to other factors being recruited to begin to repress gene classes in certain whorls as shown in Figure 7.3 (O'maoileidigh *et al.,* 2014).

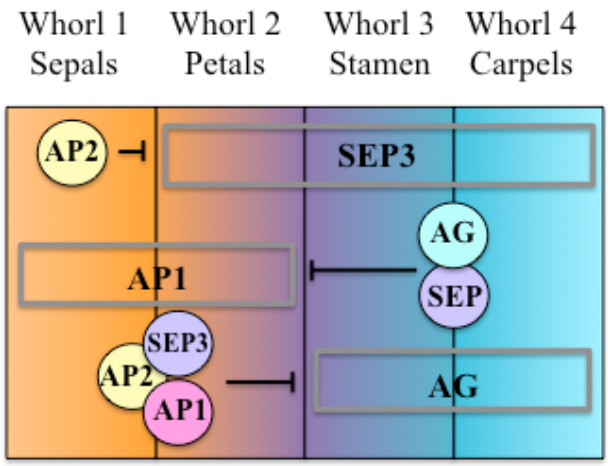

Figure 7.3. Patterns of gene induction and repression in the floral meristem that lead to the development of each distinct whorl of floral organs. Grey boxes represent promoters of the genes being repressed. The zone covered by the grey box indicates the area where expression occurs. Circles represent proteins being produced in whorls, which repress expression of other floral organ determination genes.

7.1.1.3 Proposed role of CDFs in the long day flowering pathway

In *Arabidopsis*, transgenic lines constitutively overexpressing *CDF* have been shown to delay flowering due to *CO* repression, whilst mutants of multiple *cdf*s have led to early flowering because of increased *CO* expression (Fornara *et al.,* 2009). CDF1 and 2 protein level is regulated by GI and FLAVIN-BINDING KELCH REPEAT F-BOX PROTEIN (FKF1). Blue light is required to stabilise the GI-FKF1 interaction, and so longer photoperiods lead to increased amounts of the complex, leading to increased degradation of CDF1 and CDF2, thereby mediating floral induction (Imaizumi *et al.,* 2005; Fornara *et al.,* 2009). ZTL also interacts with GI, and is able to target CDF2 for degradation in a blue-light dependent manner (Song *et al.,* 2014)

It has been shown these CDFs in *Arabidopsis* act redundantly as a mutation in one CDF can have little effect on flowering time. This redundancy has also been shown to be the case in tomato, which are day neutral plants (Corrales *et al.,* 2014).

In *Arabidopsis* the mRNA profile of three CDFs: 2, 3 and 5 were found to be high at the beginning of the light period and decrease during the day, with low level expression between 4h and 20h in 12:12 light/dark conditions (Fornara *et al.,* 2009). This was also the case in tomato, where *CDF1,2,3* and *5* all peaked at the very start of the light period (Corrales *et al.,* 2014).

CDF1 is the only DOF factor so far that has been shown to directly bind to the promoter of *CO*, and mediate its repression. As other CDF DOFs have been shown to be redundant with CDF1 through genetic approaches, it remains possible that other CDFs also bind to the *CO* promoter.

When *CDF2* was overexpressed in Arabidopsis, it caused a large reduction in *CO,* and also abolished *FT* expression (Fornara *et al.,* 2009). This further demonstrated that this family of transcription factors act to repress flowering. This supported the proposal that these DOF TFs play an antagonistic role to GI in modulating *CO* (Fornara *et al.,* 2009; Fornara *et al.,* 2015).

DOF factors have also been shown to be involved in red and far-red light signalling. For example, *COGWHEEL1* (*COG1*), is CDF-related TF which has been demonstrated to act as a negative regulator of both the phytochrome A and B (phyA and B) pathways (Park *et al.,* 2003). In Arabidopsis, it was shown that overexpression of *COG1* led to hyposensitivity to red and far-red light, whereas reduction in level led to hypersensitivity (Park *et al.,* 2003). OBF4 BINDING PROTEIN3 (OBP3) is another DOF that integrates both far-red signalling from PHYB, and blue light signalling by CRYPTOCHROME1 (CRY1), to influence hypocotyl elongation and cotyledon expansion respectively (Kang *et al.,* 2000).

7.1.1.4 Short day flowering

Whilst the majority of flowering time research has been focused on Arabidopsis which flowers in response to long days, more recently, researchers have used rice as a model for short day flowering plants.

From comparisons between the day-length sensing pathways operating in the two species, it appears that *FT* is highly conserved, but is regulated differently. In rice the homolog for *FT* is *HEADING DATE3 (Hd3)*, which is induced via the *CO* homolog *Hd1*. The mechanism for flower induction in rice therefore consists of Hd1 repressing *Hd3* in long days, but under short days, Hd1 induces *Hd3* expression, leading to flowering (Izawa *et al.,* 2002). *Hd1* and *CO* show the same expression pattern, and appear to both be regulated by the circadian clock. The rice GI homolog (OsGI) has also been shown to regulate *Hd1*, just as GI does with *CO* in *Arabidopsis* (Hayama *et al.,* 2003).

Research into flowering time has also demonstrated that whilst LDPs appear to find the length of the light period more important, SDPs are usually more dependent on a long dark period. SDPs are more sensitive to night break treatment, where impulses of light are introduced during the dark period, but are usually less sensitive to light quality during the light, in relation to flowering (Park *et al.,* 2003). Most LDPs are more sensitive to light quality and many require far-red light at the end of the long photoperiod in order for a long day to be interpreted as a signal leading to the induction of flowering.

Interestingly, flowering pathway components with no homologs in *Arabidopsis* have been identified in rice. One example is *EARLY HEADING DATE1* (*Ehd1*), which has been shown to promote flowering preferentially under short days, even without functional Hd1 (Doi *et al.,* 2004). Analysis suggests this gene functions upstream of

Hd3 (*FT*). It was shown to peak before dawn, and only displayed a high level of induction under short days (Wu *et al.,* 2008).

In other SDPs, such as the Japanese Morning Glory (*Pharbitis nil)*, it was found that the two FT homologs: *PnFT1* and *PnFT2* were only induced if the dark period was long enough (Hayama *et al.,* 2003). Furthermore in these plants depending on light regime *PnFT* and *PnCO* expression could be uncoupled, without effects on flowering. This therefore suggests that an additional pathway may exist for *FT* induction. This could mean SD plants have extra flowering induction pathways yet to be discovered.

7.2 Results

7.2.1 Impacts of changes in *KfCDF2* expression levels on the developmental progression into flowering

Despite *K. fedtschenkoi* being a SDP, *KfCDF2_FL* overexpressing transgenic lines were found to flower under 16:8 light: dark long-day conditions (LDs). The *KfCDF2* over-expresser lines were observed to transition to flowering at various times of the year, suggesting that the influence of *KfCDF2* over-expression on the induction of flowering was independent of the annual changes in day-length and light quality and quantity falling naturally on the greenhouse. Long-days (16:8) were maintained in the greenhouse using supplementary lighting, which maintained 16:8 conditions throughout the year.

It was therefore important to investigate this constitutive flowering phenotype in the *KfCDF2* over-expresser lines and to try to identify components of the molecular mechanism under-pinning this photoperiod independent flowering. These initial findings also suggested it was important to investigate both the over-expression and RNAi knockdown lines for *KfCDF2* under a variety of light regimes to gain further insights into the impact of *KfCDF2* levels on flowering and the speed of floral development.

Three pots, each containing three clonal plants of each line: wild type, *KfCDF2_FL_14C* or *KfCDF2_RNAi_19A,* were placed in three different Snijders Microclima MC-1000 growth cabinets under different light/ dark regimes. One cabinet was maintained under 16:8 LD (long day) conditions, which does not induce flowering in the wild type. A second cabintet was maintained under 12:12 LD, which can induce flowering if plants are maintained under LD conditions for several months,

and the third cabinet was maintained under 8:16 LD (short day) conditions, which induces flowering in *K. fedtschenkoi*.

Plants were grown for 80-days under these conditions. Figure 7.4 shows one pot of each plant type as they entered each condition, and Figure 7.5 shows the same pot at the end of the 80 days.

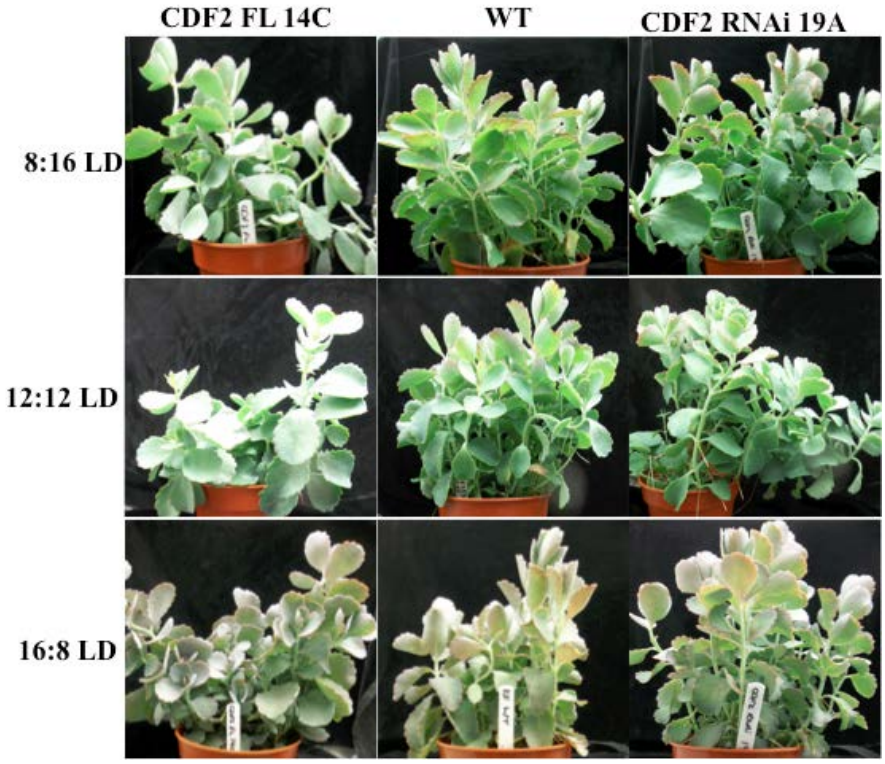

Figure 7.4. Starting phenotypes of *K. fedtschenkoi KfCDF2* transgenic lines subjected to different day-length regimes. *K. fedtschenkoi* wild type, *KfCDF2_FL_14C* and *KfCDF2_RNAi_19A* were grown in 16:8 LD (long day, non-flowering conditions) initially for 6 months. Plants were then placed into different growth cabinets under the different light regimes indicated on the left hand side of the figure.

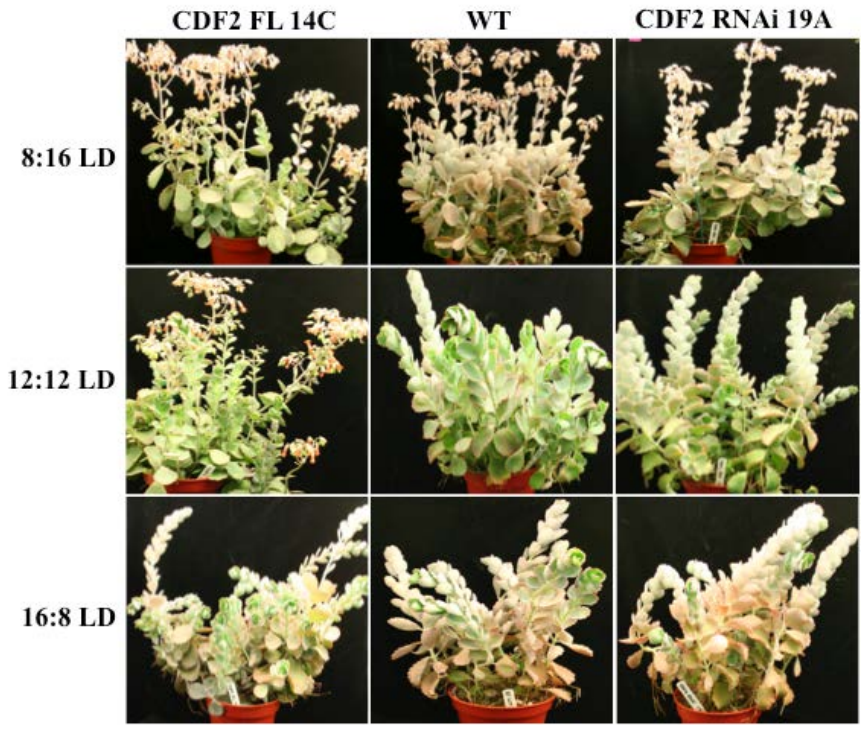

Figure 7.5. Progression of floral induction under different day-length regimes in wild type and *KfCDF2* over-expressor and RNAi liens of *K. fedtschenkoi*. *K. fedtschenkoi* wild type, *KfCDF2_FL_14C* and *KfCDF2_RNAi_19A* were grown under the indicated day-length regimes for 80-days in total. In 8:16 LD all three lines flowered, but *KfCDF2_FL_14C*s first flower opened 17-days earlier than the wild type and *KfCDF2_RNAi_19A*. In 12:12 LD conditions, *KfCDF2_FL_14C* also flowered only 3 days slower than in the 8:16 LD condition, whereas *KfCDF2_RNAi_19A* only showed a hooked stem, and wild type had yet to form a hook. In 16:8 no lines had hooked their meristematic region, but *KfCDF2_FL_14C* had started to lengthen its stems and was producing smaller leaves, which are both characteristic of the progression towared meristem hooking and floral initiation in *K. fedtschenkoi*.

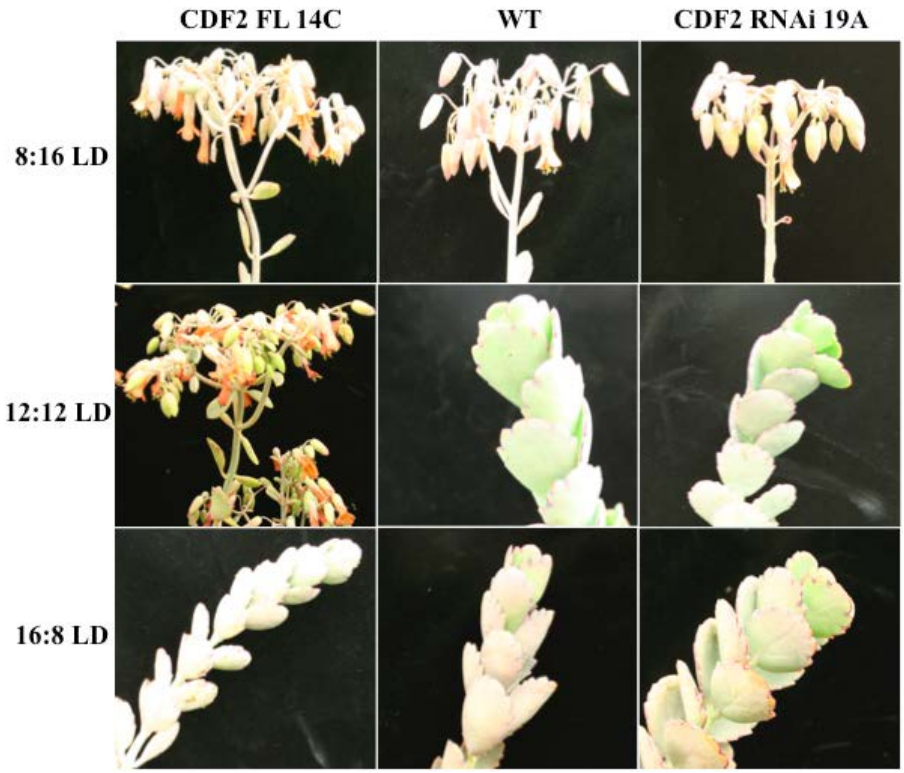

Figure 7.6. Close-up images of a stem from each of the *K. fedtschenkoi* lines at the end of 80-days in various light regimes. Light regimes are listed on the left-hand side and line names are listed above. By the end of the experiment, *KfCDF2_FL_14C* plant were producing many flowers in both 8:16 (short days) and 12:12 conditions. For Wild Type and *KfCDF2_RNAi_19A,* there was only one open flower, whilst the rest were still unopened buds under 8:16 SD conditions. For *KfCDF2_RNAi_19A* in 12:12 LD, the hooking-over of the mersitem could be seen. In 16:8 LD (long days) there was no visible hooking or flowering for any of the lines, but the stem of *KfCDF2_FL_14C* had elongated and the newly developed leaves on the elongated stem were smaller than those on the other two lines.

In 8:16 conditions (SDs), which induced flowering in all lines, the over-expressing line flowered 17-days earlier than the wild type and RNAi line. Flowering was classed as the day the first flower became visible and was fully open. Also, in short days, the *KfCDF2_FL_14C* plants had, on average, more open flowers (73) after the 80-day observation period, compared to wild type (3), or *KfCDF2_RNAi_19A* (2). Furthermore, in 12:12 LD there was a much more striking difference, as the wild types had yet to begin to hook after 80-days, whilst the *KfCDF2_FL* had produced 68 flowers (Fig. 7.5; Table 7.1). From the 12:12 results though, it appears that the RNAi

lines may also show a slightly more rapid flowering response than wild type due to visible stem hooking (Fig. 7.6). In long-days (16:8 LD), no hooking or flowering occurred, but it may take longer than 80-days to achieve the required threshold level of a floral induction factor in the *KfCDF2_FL_14C* line which had been observed to develop flowers under 16:8 LD conditions in the greenhouse. In the over-expressing line under the 16:8 long day conditions, stems had begun to elongate, and leaves had become smaller, characteristic of the start of flowering (Fig. 7.6). Furthermore, in the greenhouse, plants that had been kept under long days (16:8 light/dark conditions) for over 6-months were also not yet flowering, which may suggest that plant age plays an important role in flowering initiation in the *KfCDF2_FL* over-expresser lines.

Table 7.1. The influence of different day-length regimes on various traits indicative of the developmental progression towards flowering in *K. fedtschenkoi.*

	Days taken in 8:16	Days taken in12:12	Days taken in16:8
FL Hook	23	35	-
FL Flower	63	66	-
WT Hook	35	-	-
WT Flower	80	-	-
RNAi Hook	36	80	-
RNAi Flower	80	-	-

The number of days required to achieve stem hooking and the opening of the first flower under each light regime for each line tested. The experiment terminated after 80-days, and so dashes indicate that the line did not reach that flowering stage within the time of this experiment.

Under 12:12 light/ dark conditions the biggest differences in flowering induction was observed, but due to time constraints, and because all lines flowered under 8:16 light/ dark cycles, the gene expression investigation on the genetic control of flowering time was conducted using the samples collected from the 8:16 light/ dark conditions, so that the developmental progression towards flowering in all three lines could be observed and compared.

7.2.2 Gene expression changes in response to a flowering stimulus

To investigate gene expression changes involved in flowering time regulation in the *KfCDF2* lines, a time course was set up. Six-month old plants of each genotype were raised in non-inductive conditions (16:8 LD), and week 0 samples were collected. For all remaining weeks of the time course, the plants were under short days (8:16 LD). Wild type lines successfully flowered in 8:16 LD conditions after 84 days (12 weeks). Samples of a range of above-ground organs and tissues were collected each week to allow the investigation of flowering pathway gene transcript levels in various parts of the plant. The samples collected included LP1, 3, 6 and 10, stem and shoot apical meristem (SAM). As plants progressed to flowering, the hooked region of the stem at the growing point, the developing buds and flowers were also collected.

Various *K. fedtschenkoi* orthologs of flowering pathway genes were selected for investigation using RT-qPCR. The flowering genes were chosen based on their known roles in the regulation of flowering time in Arabidopsis. The Arabidopsis homologs were identified in the draft genome and assembled transcriptome of *K. fedtschenkoi,* and RT-qPCR primers were designed for their amplification and quantification. A gene referred to as *Contig9471* was used as the reference gene. This gene in *K. fedtschenkoi* has highest homology to At2g30720, which encodes a thioesterase found in the mitochondrion of cells. It was found to show very little change in expression level throughout the 24 h cycle, and so was used as the reference gene for the relative quantification of gene transcript abundance levels for the flowering pathway genes.

Table 7.2. Identified flowering genes in *K. fedtschenkoi* with accession numbers in Arabidopsis and Chrysanthemum.

Gene Name	Symbol	*Arabidopsis* Accession Number	Number of hits in *K.f.*	*C. seticuspe f. boreale* Accession Number	*Kf* Genome accession numbers	*Kf* gene symbol
CYCLING DOF FACTOR2	*CDF2*	At5g39660	1		KF17950	*KfCDF2*
GIGANTEA	*GI*	At1g22770	1		KF96760	*KfGI*
CONSTANS	*CO*	At5g15840	2		KF105990	*KfCO25*
					KF127450	*KfCO46*
ZEITLUPE	*ZTL*	At5g57360	1		KF118840	*KfZTL*
SUPRESSOR OF OVEREXPRESSION OF CONSTANS1	*SOC1*	At2g45660	1		KF03655	*KfSOC1*
SEPALLATA	*SEP2*	At3g02310	1		KF22770	*KfSEP*
	SEP3	At1g24260				
FRUITFUL	*FUL*	At5g60910	1	AB839766	KF111895	*KFFUL*
APETALA1	*AP1*	At1g69120	1	AB839766	KF135520	*KfAP1*
FD	*FD*	At4g35900	1	AB839768 /9	KF106650	*KfFD*
FLOWERING LOCUS T	*FT*	At5g03840 At1g65480	2	AB839766 AB839767	KF67365 ----	*KfFT12* *KfFT18*

Genes involved in flowering time that were identified in Arabidopsis (TAIR entries) and *Chrysanthemum seticuspe f. boreale* (DDBJ entries)(Higuchi *et al.,* 2013) and orthologs that were indentified within the *K. fedtschenkoi* transcriptome and genome. *KfFT18* could not be found in the draft assembly of the *K. fedtschenkoi* genome, but was found in the assembled transcriptome as Contig18180.

7.2.2.1 Wild type progression to flowering

Initial samples were collected on day 0 from plants that had been grown under long-day (16:8 LD) conditions since they were small, developmentally-synchronised adventitious leaf-margin plantlets. The light regime was changed to short days (8:16 LD). Sampling was then conducted after various numbers of weeks: 1, 5, 7, 9, 11 and 12. After 5-weeks, the plants had developed a noticeably hooked over stem at their growing point. By week 7, small buds had begun to form, and by week 11 large buds were forming. In the 12th week, flowering occurred in wild type.

In older leaves, *KfCDF2* was an abundant transcript at week 0, consistent with its original identification as putative CAM-associated, clock-controlled TF (see Chapter 3). However, in all other tissues, *KfCDF2* transcripts were low until after week 5 of SD conditions (Fig. 7.7). *KfCDF2* was thereafter present in all tissues at a high transcript abundance until buds began to develop, at which point *KfCDF2* transcript levels declined. In the meristems, dark levels of *KfCDF2* increased between week 5, when the stems began to hook, and week 11, when bud development had been established. These results for wild type plants suggested that *KfCDF2* plays a role in promoting the developmental progression to flowering, but is of lesser importance during floral development in the meristematic tissues.

As *KfGI* is a core circadian clock gene, it is not surprising that its transcript levels remained relatively constant in the leaves throughout the short-day inductive timecourse (Fig. 7.8). After 7-days under short days, *KfGI* increased in transcript abundance in the light in all leaf pairs (week 1), which is likely sending an inductive signal to the flowering pathway. *KfGI* transcript levels subsequently began to increase in the light in the stem, up to week 9, when visible buds had developed. Levels of *KfGI* in the meristem itself did not change dramatically over the duration of the experiment, although the light peak of *KfGI* did show a general decline as the weeks under short day conditions progressed. During bud development, the level of *KfGI* increased, but subsequently dropped when the flowers opened, suggesting *KfGI* may play a role in regulating floral development, or have some role in floral development itself, in addition to its role in linking the clock and flowering time pathways.

For *CO*, two partial sequences were found in the *K. fedtschenkoi* transcriptome, which show high sequence homology: *KfCO25* and *KfCO46*, and so both transcripts were investigated (Fig 7.9 & Fig. 7.10). Their nucleotide sequences show 84.5 % identical bases, and at protein level, their predicted amino acid sequences shared 83.5% pairwise identify. Both genes showed very similar results, except that *KfCO46* was a lower abundance transcript, and also appeared to show a less robust oscillation between the light and the dark sample.

Both *KfCO* genes were not detected at either the light or dark timepoints samples in this experiment before week 5. In week 5, both *KfCO* transcripts were detected in leaves, and their transcript abundance then increased steadily increases. In leaves and the meristem, *KfCO* transcript abundance peaked in week 11, whilst in the stem it peaked at week 12. *KfCO25* transcript levels declined steadily in buds as they developed to flowers, whereas *KfCO46* levels remained high, in both light and dark, until week 12, when flowers opened. Based on these results, *KfCO46* could play a role in floral determination, whereas *KfCO25* may play a larger role in the progression from the vegetative meristem to the inflorescence meristem.

For FT, two genes were detected in the *K. fedtschenkoi* genome: *KfFT12* and *KfFT18*. Amino acid sequence alignments indicated that *KfFT12* shared highest sequence homology with antiflorigen (At5g03840), whereas *KfFT18* shared greatest homology with florigen (At1g65480). However, the RT-qPCR results presented here indicate that *KfFT18* may act as the antiflorigen in *K. fedtschenkoi* (Fig. 7.11 & Fig. 7.12). *KfFT18* transcript levels were found to be high in non-flowering conditions, after which levels began to decrease, especially in the light period, where this gene had displayed higher transcript levels relative to the dark. *KfFT18* was barely detected in

buds or flowers, except at week 12, when it was detected primarily in the dark period. *KfFT12* on the other hand had low transcript levels during non-inductive flowering conditions, and levels began to increase steadily once the plants were placed into conditions that promoted flowering (short days). *KfFT12* also peaked during the dark rather than the light, and was found to be present at very high transcript levels in buds and flowers. FT is thought to be a key component of the mobile flowering integrator signal, which is able to travel from leaves into the stem up to the meristem to mediate the induction of flowering.

For *KfAP1*, which is important not only for progression to flowering, but also in floral development, the transcript levels were high throughout most of the samples and weeks (Fig. 7.13). In LP10 once entering into flowering conditions, expression levels increased each week. In the meristem, levels of *KfAP1* were higher from week 7 onwards, and transcript levels were also shown to increase in buds as they develop, peaking in well-developed buds and flowers at week 12. Interestingly, in the stem there appeared to be a large final signal sent via the stem at the end of week 12, coincident with the flowers opening, which may suggest a role for this gene in the induction of flower opening.

With *KfSEP,* it was detected in all tissues at all times, except for the light period of week 0 (in long days; Fig. 7.14). In all tissues there appeared to be repression in light in non-flowering conditions (long days). Once plants entered short days, this repression seemed to be removed (Fig. 7.14). In some weeks, *KfSEP* transcript levels were higher in the light, whilst in other weeks transcript levels were higher in the dark, suggesting this gene may not be under circadian control in *K. fedtschenkoi*, but

this may be a result of phase changes in the timing of peak transcript abundance which could only be characterised with a more detailed timecourse with samples collected at more regular intervals over the light/ dark cycle. From all the tissues and timepoints measured, *KfSEP* transcript levels were at their highest in the meristems and buds, especially when the buds were large and beginning to open (week 12). These findings are consistent with the known role of *SEP* in co-regulating floral development all flowering whorls, which suggests it may also be involved in all stages of floral development.

KfZTL is another gene that functions within the clock, where it has specifically been demonstrated to function in targeting TOC1 ubiquitination and degradation via the 26S proteasome. ZTL has also been shown in Arabidopsis to mediate the degradation of CDF protein, especially CDF2 (Fornara *et al.,* 2009). Near the end of this flowering experiment, after buds became visible, *KfZTL* levels increased in all tissues (Week 11), with the highest transcript levels detected in LP10 (Fig. 7.15). *KfZTL* also reached its peak transcript levels in floral organs when buds were beginning to open, suggesting that it could be involved in targeting other flowering pathway proteins for degradation, to enable flowers to open and prevent floral reversion.

KfSOC1 is induced to promote flowering. In older leaf pairs, (LP6 and 10) there were high levels of the transcript in the dark period in non-flowering conditions (Fig. 7.16). As progression to flowering continued, *KfSOC1* transcript levels were detected in LP1 and 3, whilst lower levels were detected in LP6 and 10, likely due to it being repressed to prevent floral reversion. Levels also began to decrease in the meristem and the stem between weeks 7 – 9, and this may possibly also be a sign that this

regulation is linked to the plant avoiding floral reversion. Furthermore, *KfSOC1* levels increased steadily in buds, and peaked at bud opening, and so *KfSOC1* may play a further role in the final flower-opening signal.

FD is thought to interact with FT and induce flowering genes. Interestingly, in *K. fedtschenkoi* the transcript levels of *KfFD* were seen to decrease once plants entered inductive conditions for flowering. *KfFD* levels began subsequently to rise again from week 6 in older leaves, but were still significantly lower than in Week 0 (Fig. 7.17). In the stem and meristem, levels of *KfFD* were high until Week 5 and then declined. *KfFD* also increased in floral organs throughout development. This may be because if *KfFT12* is moving to the meristem, then this is the only place that *KfFD* transcript levels need to be maintained at a high level to enable the FT-FD complex to form.

KfFUL, which is essential for flowering, was barely detectable in any tissue in non-flowering conditions, but once the conditions had changed to 8:16 LD, levels of this transcript began to increase until the approximate time of stem hooking (week 5; Fig. 7.18). This supports the proposal that this gene maintains a conserved function between Arabidopsis and *K. fedtschenkoi*, being involved in the transition from inflorescence to floral meristem. *KfFUL* levels also increased in buds during their development, with slightly lower levels found in flowers.

Finally, *LEAFY* (*LFY*), which is another essential gene in the pathway of flowering time regulation and floral development in Arabidopsis (Fig. 7.1), was only able to be identified in the *K. fedtchenkoi* genome (KF07545). It was also amplified successfully from genomic DNA of *K. fedtschenkoi* using the designed qPCR primers. However,

KfLFY could not be found though in the assembled transcriptome data, nor could it be amplified in pooled flowering time samples using RT-qPCR. *KfLFY*s role in flowering progression could not therefore be determined in this study, as its expression could not be confirmed, suggesting LFY may not play an active role in *K. fedtschenkoi* development, or that its expression was missed in the range of samples used in this experiment; for example its expression may be confined to a particular time of day or to specific tissues or cell types which were not well represented in the RNA samples used in these experiments.

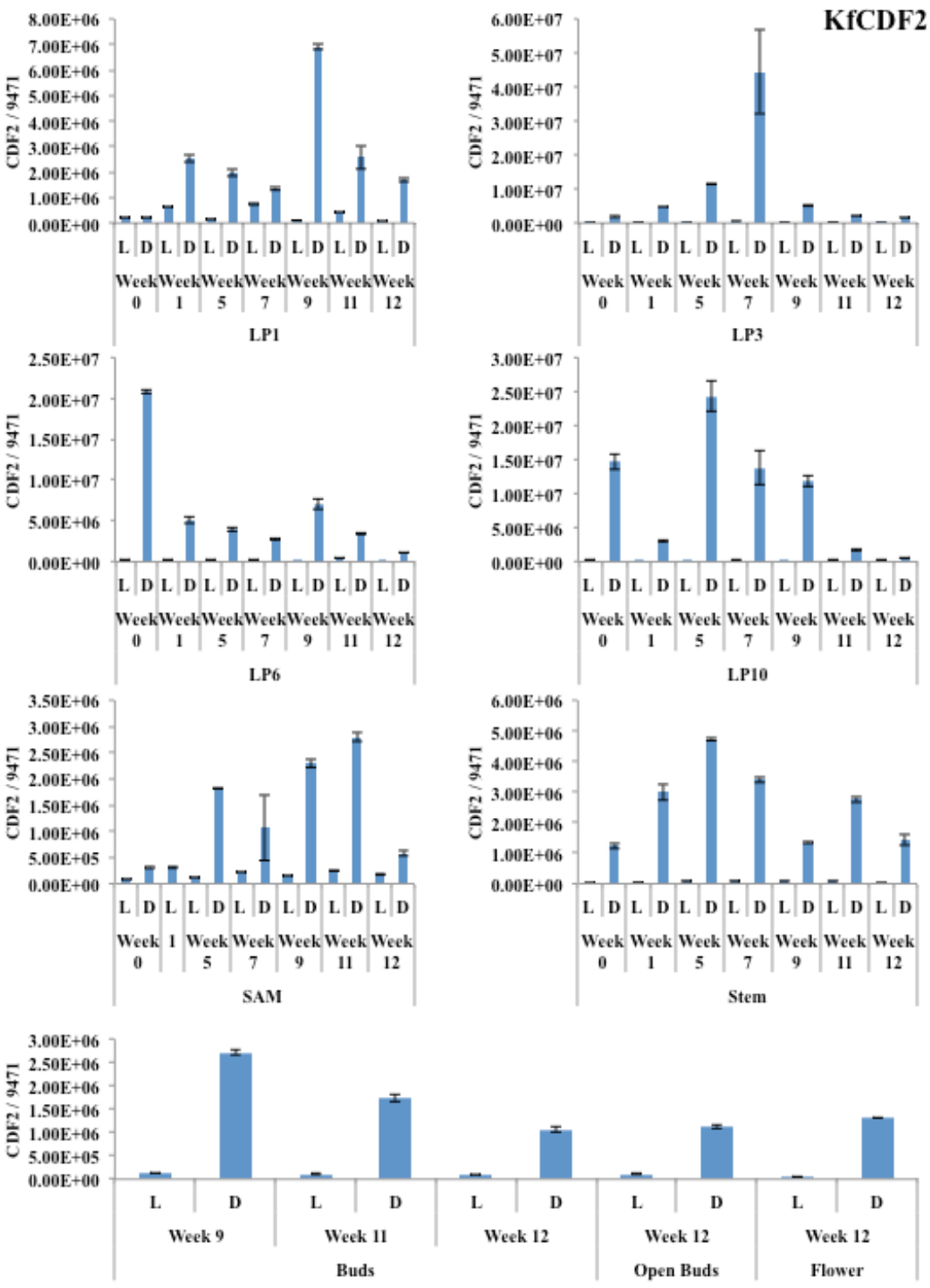

Figure 7.7. RT-qPCR data for *KfCDF2* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

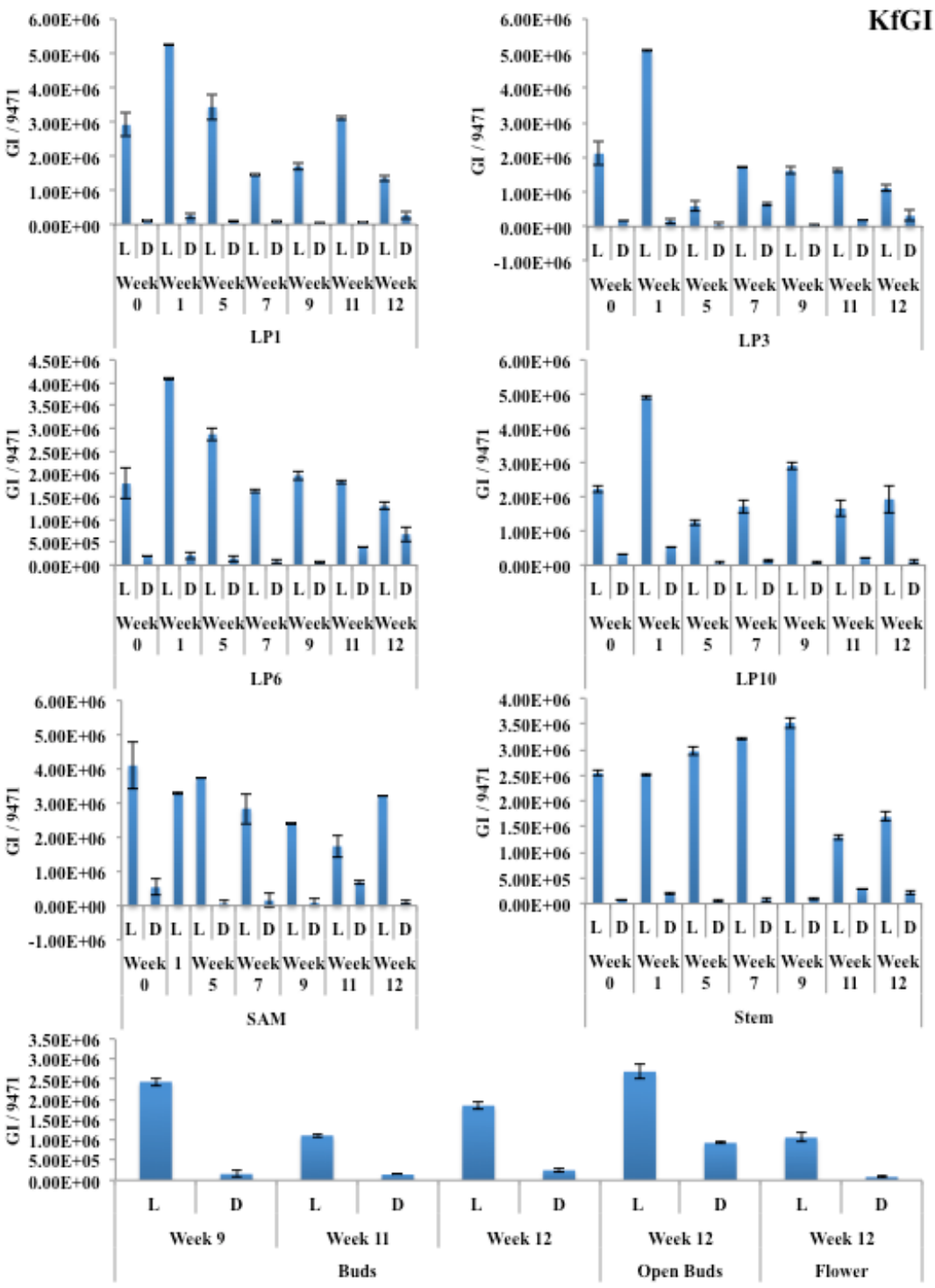

Figure 7.8. RT-qPCR data for *KfGI* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

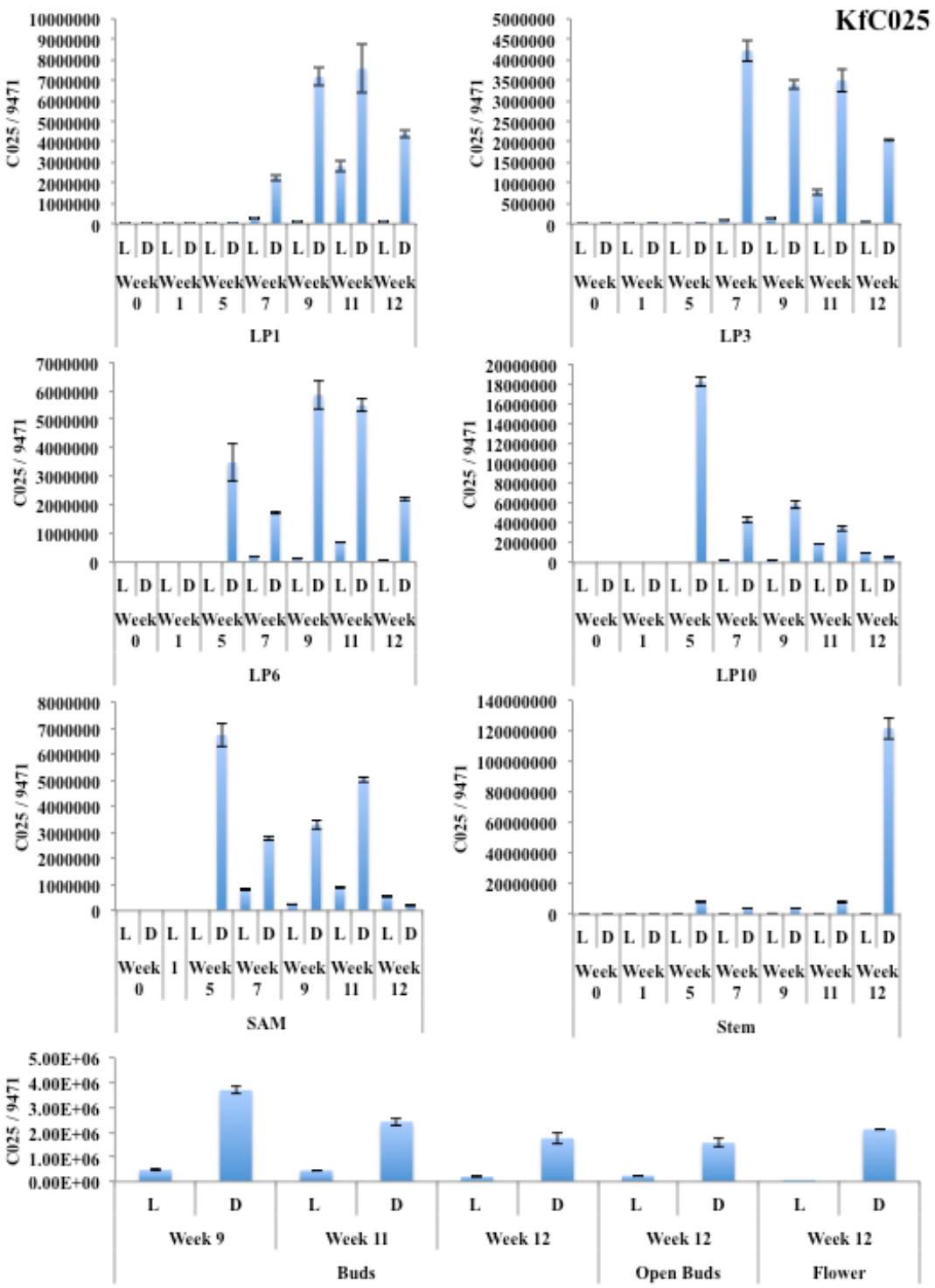

Figure 7.9. RT-qPCR data for *KfCO25* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

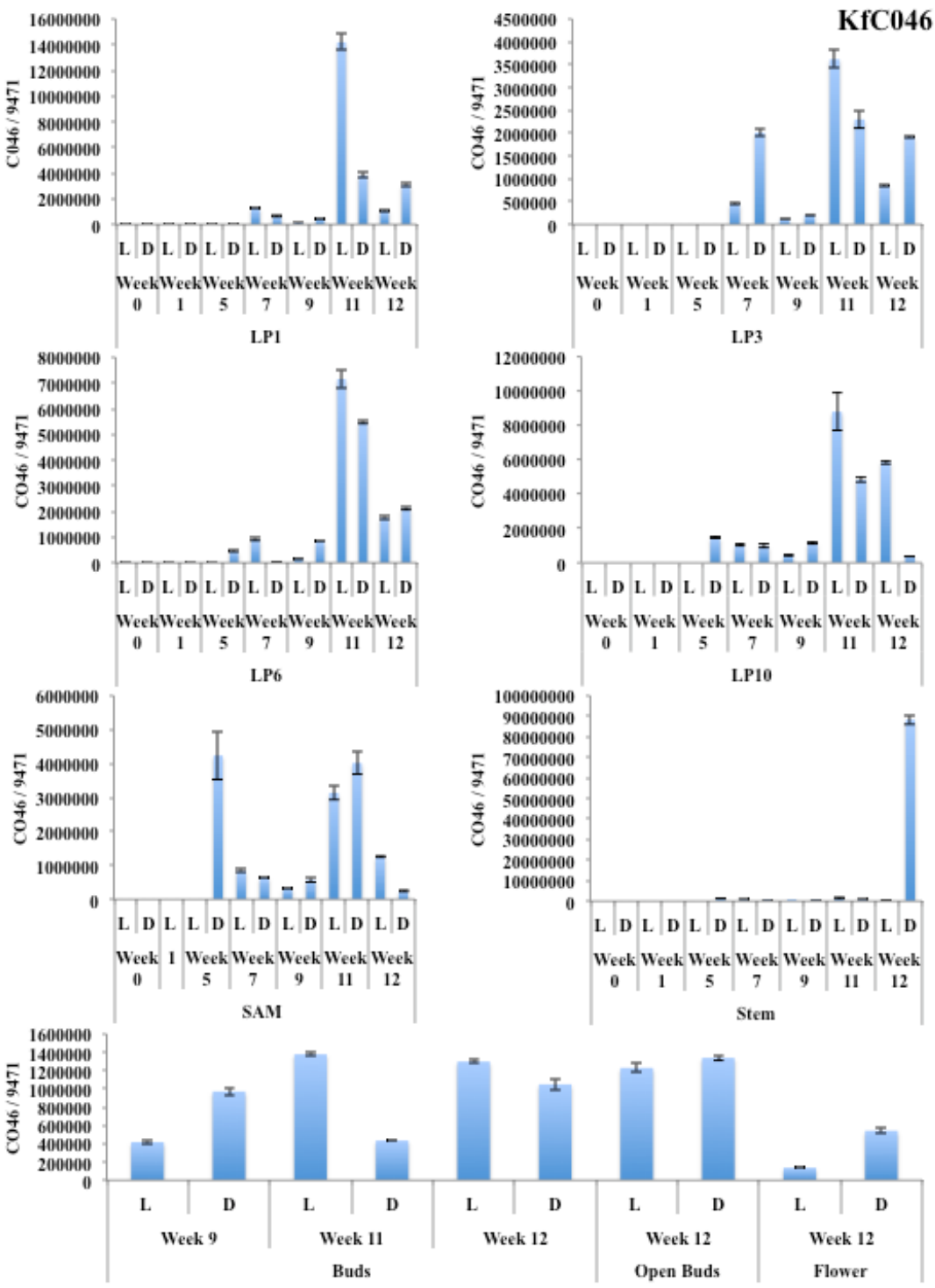

Figure 7.10. RT-qPCR data for *KfCO46* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

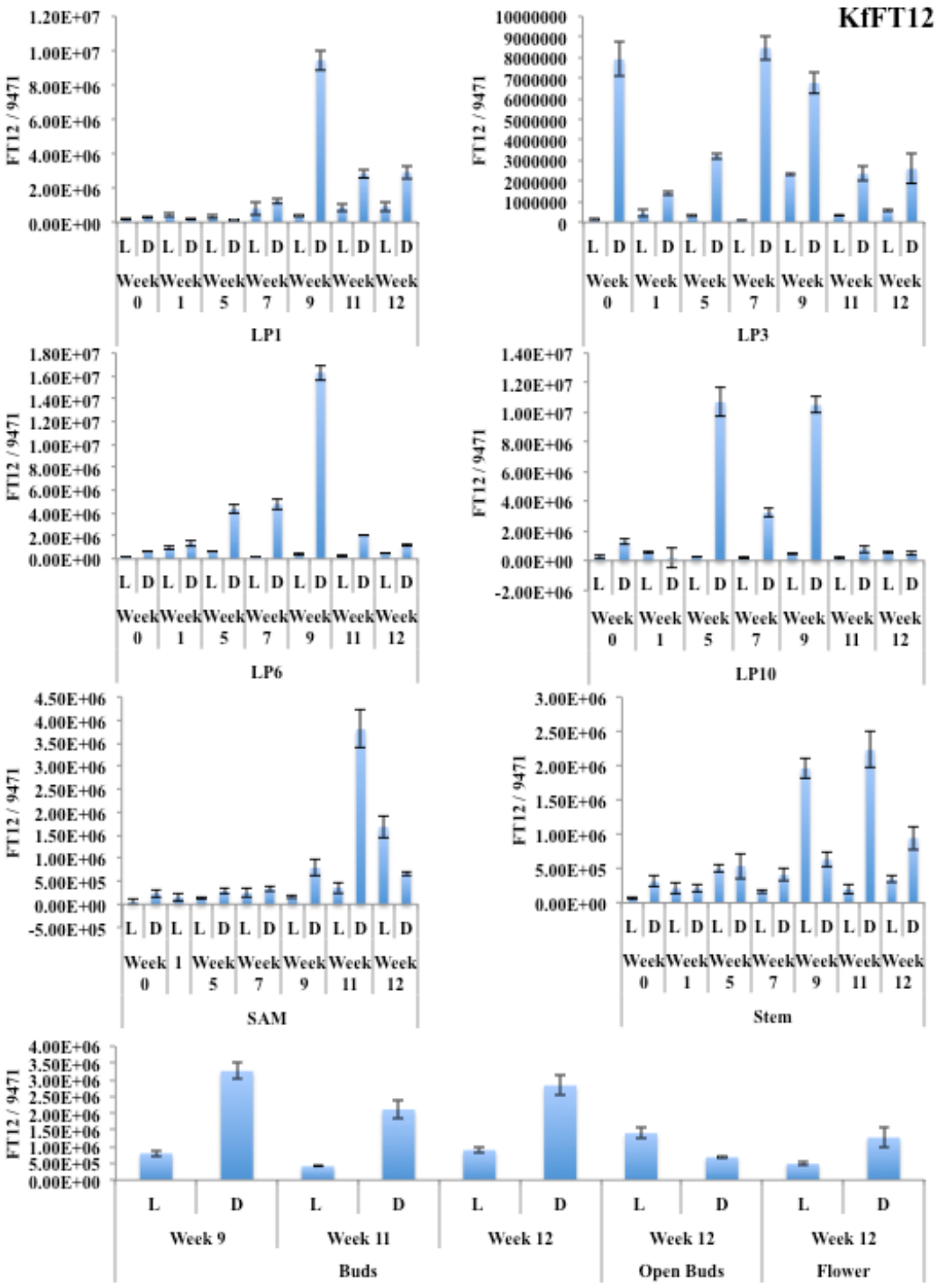

Figure 7.11. RT-qPCR data for *KfFT12* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

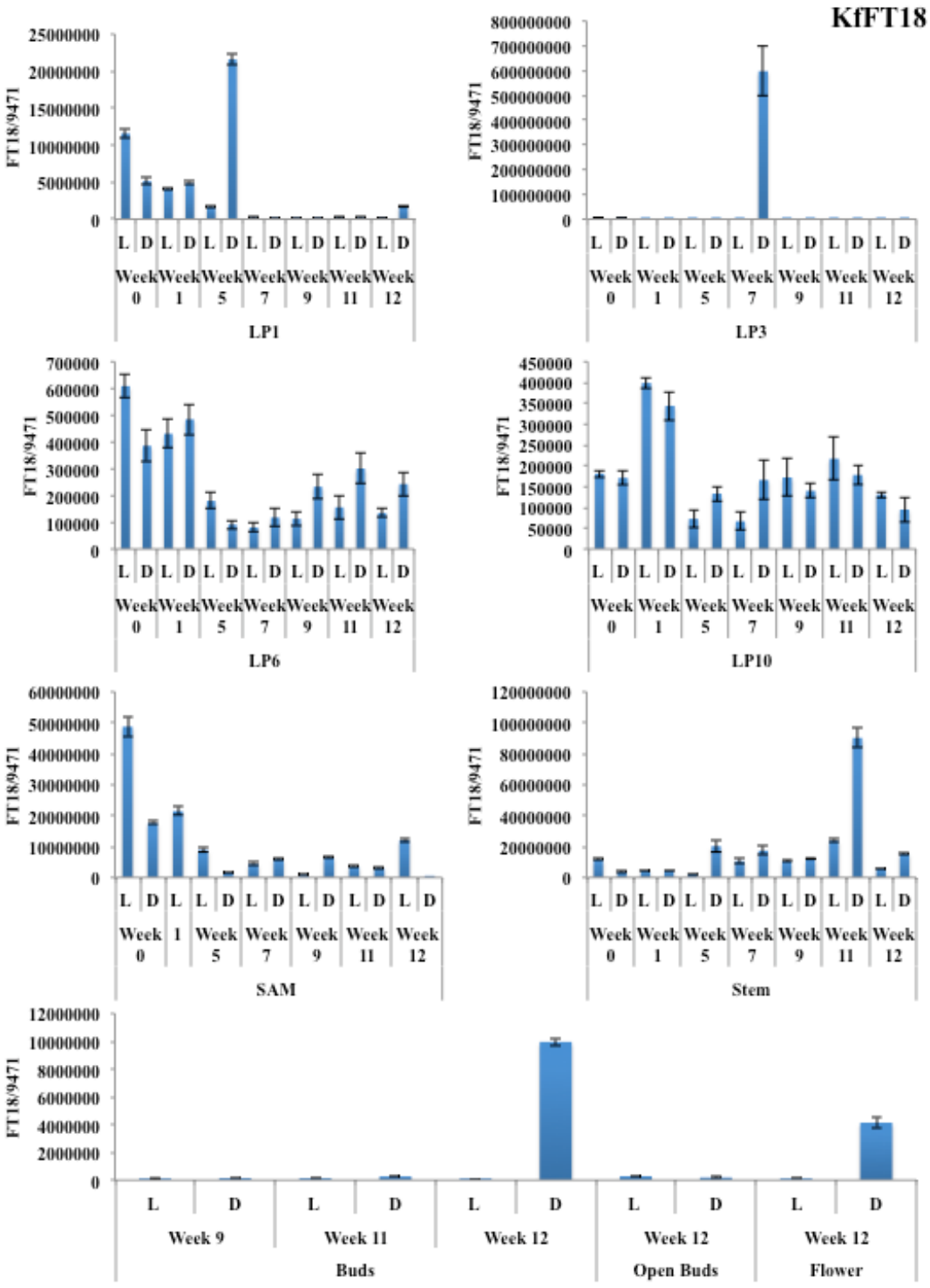

Figure 7.12. RT-qPCR data for *KfFT18* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

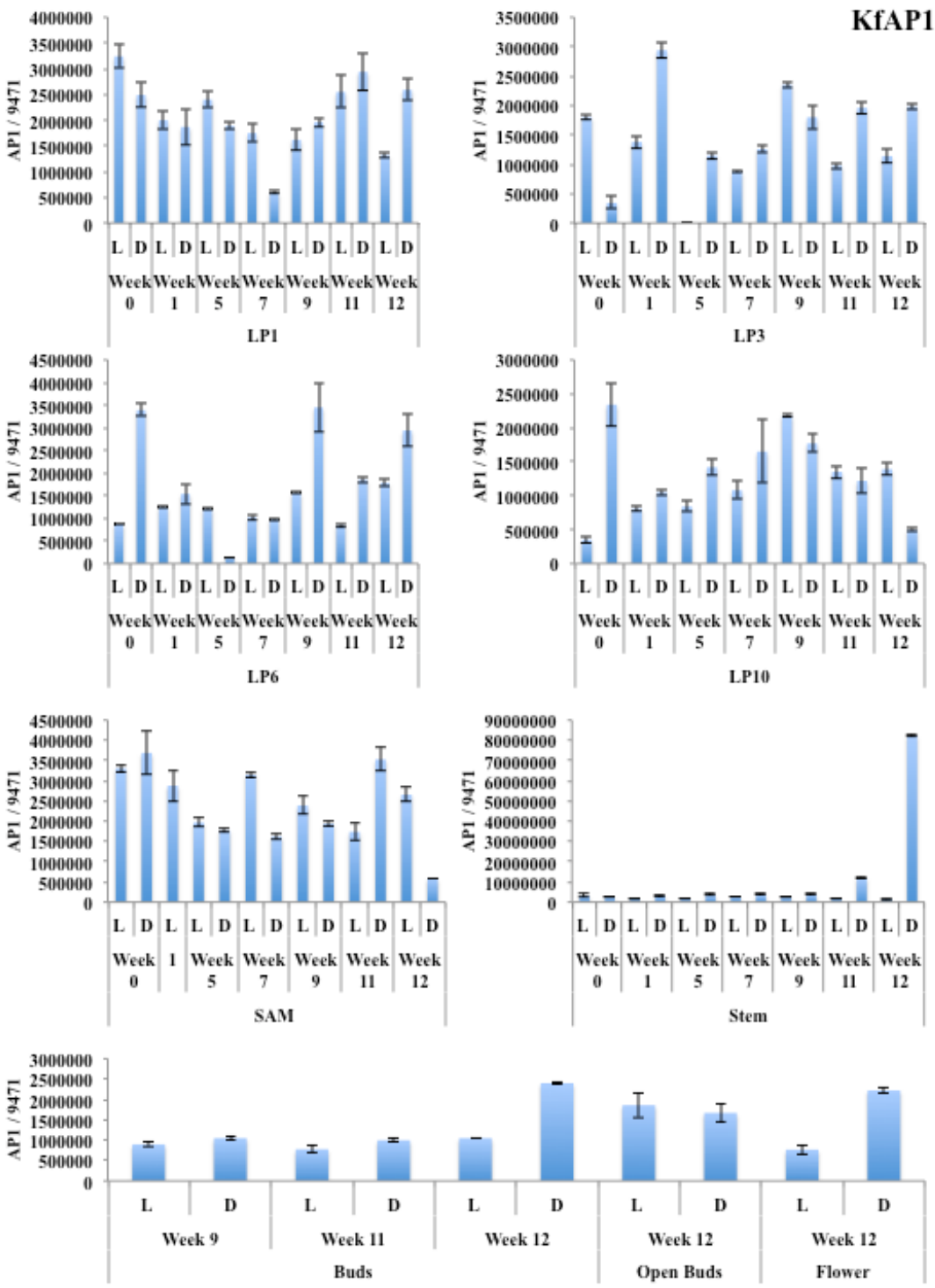

Figure 7.13. RT-qPCR data for *KfAP1* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

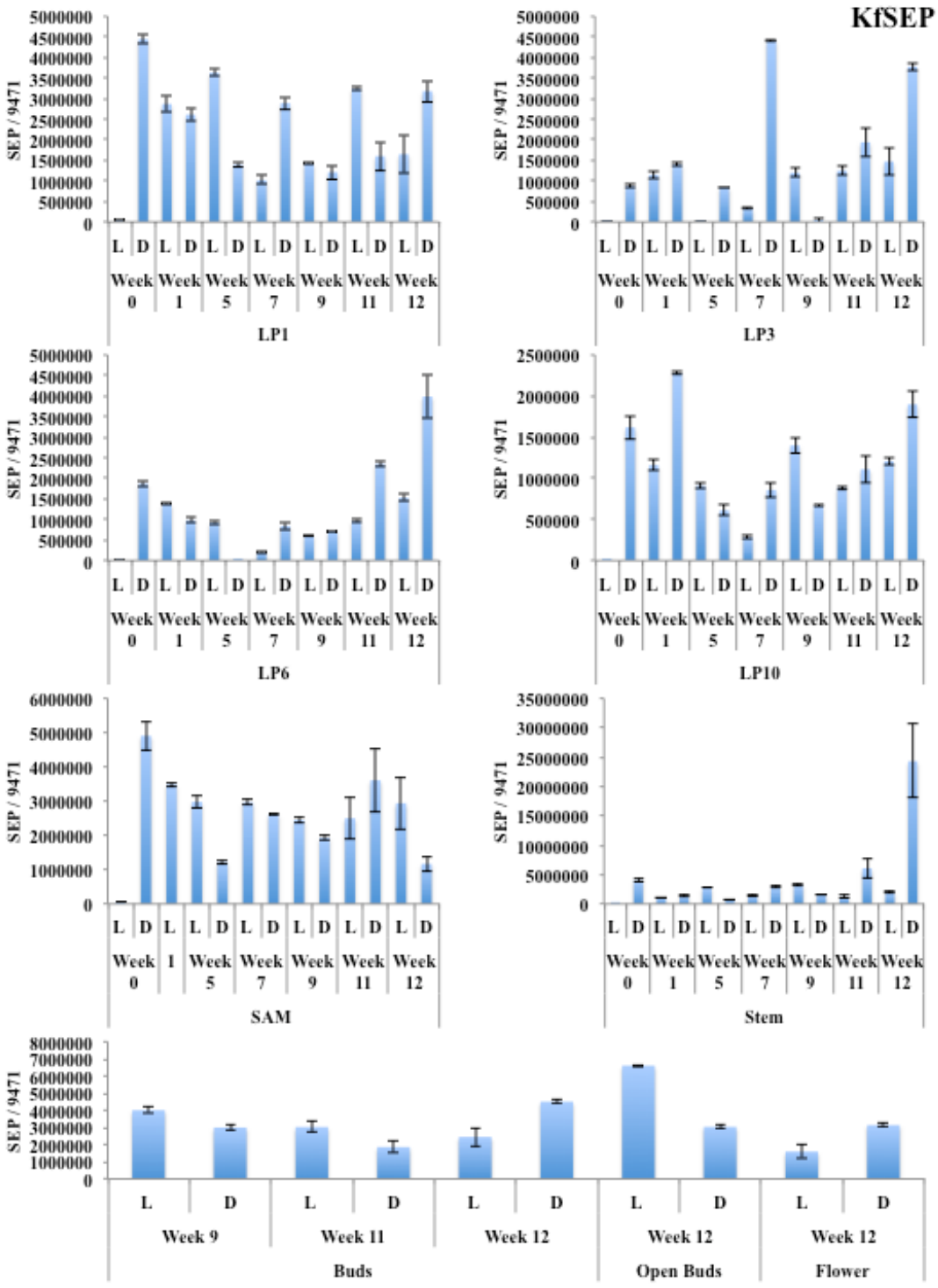

Figure 7.14. RT-qPCR data for *KfSEP* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

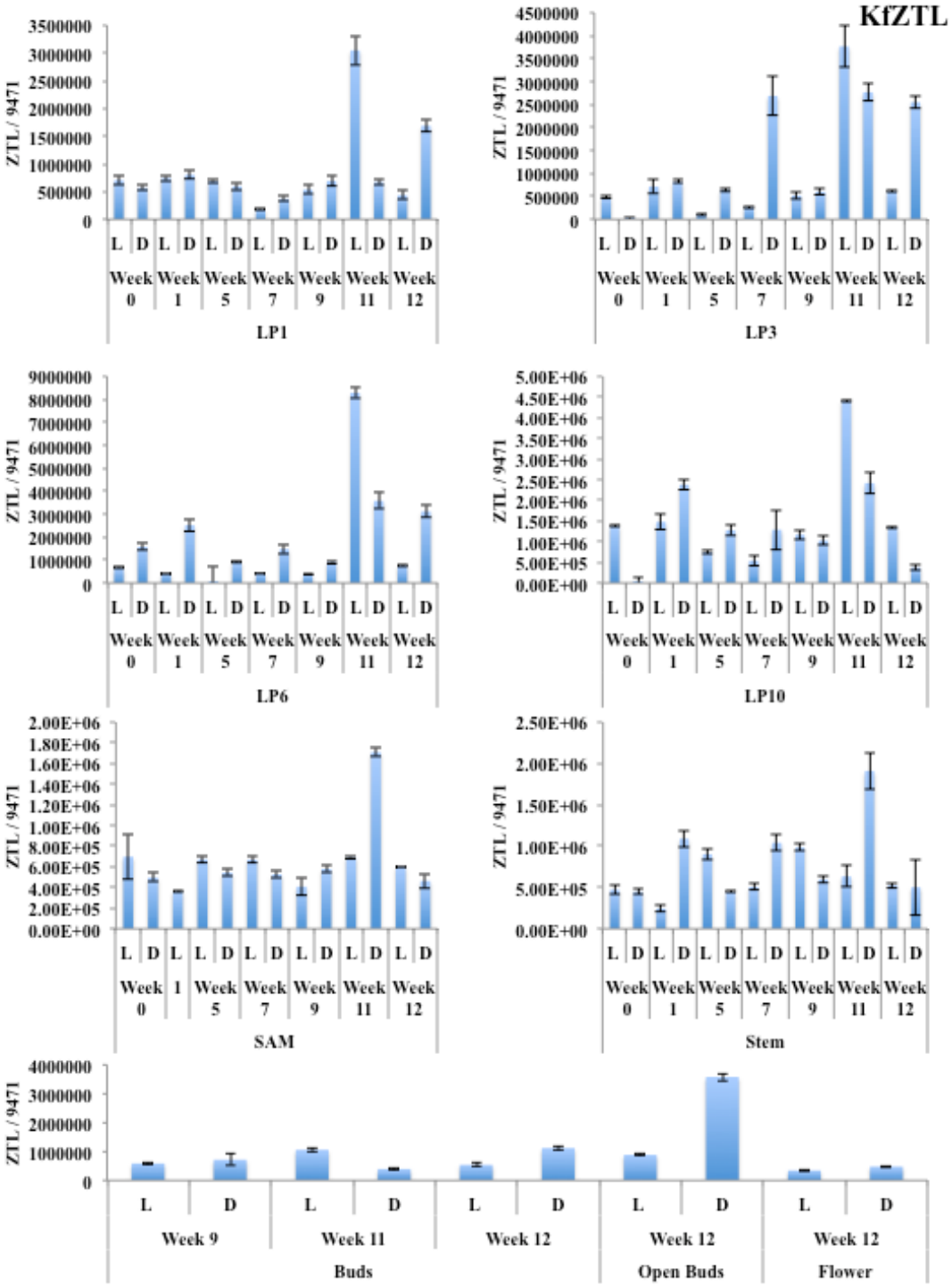

Figure 7.15. RT-qPCR data for *KfZTL* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

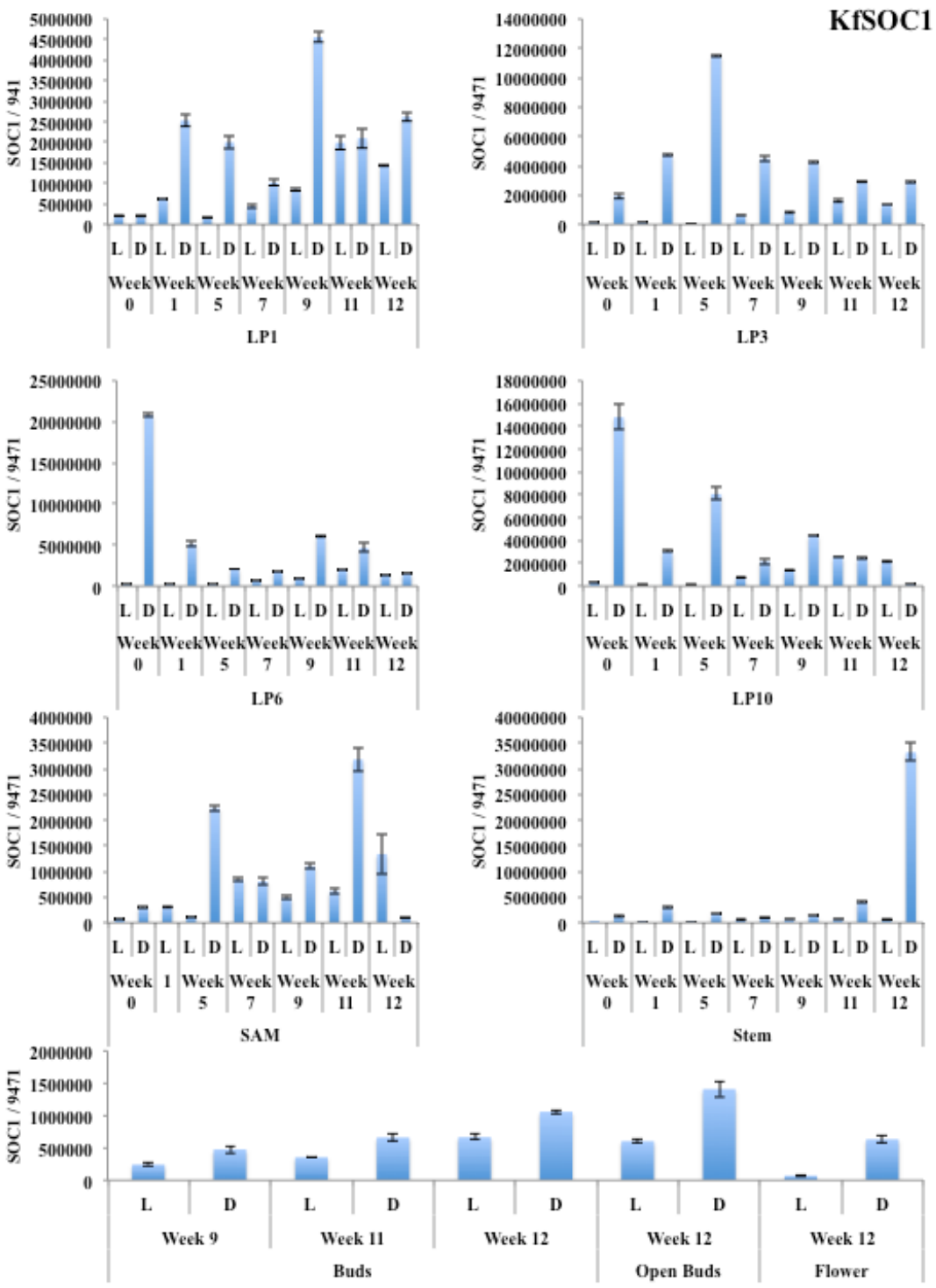

Figure 7.16 RT-qPCR data for *KfSOC1* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

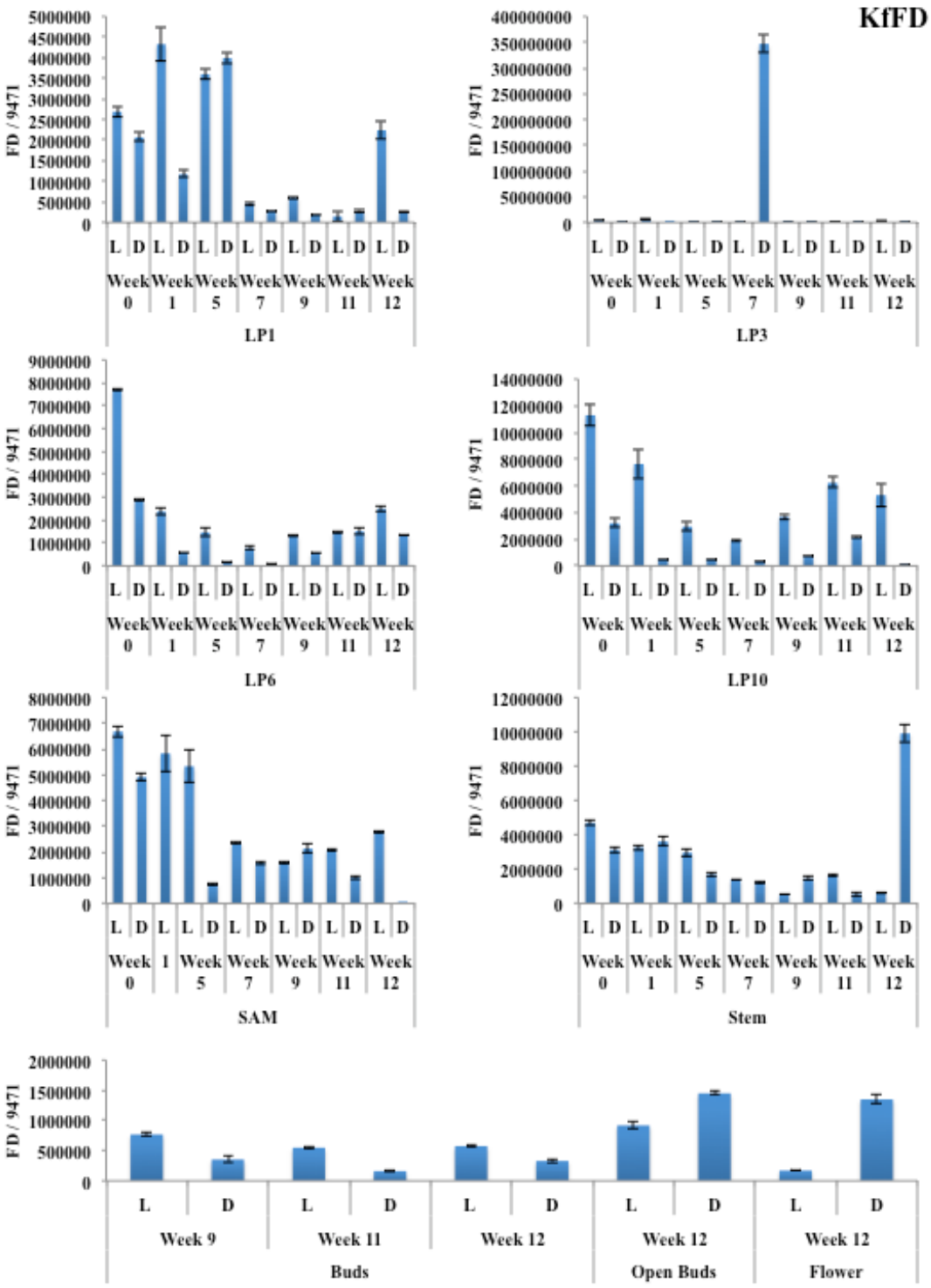

Figure 7.17. RT-qPCR data for *KfFD* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

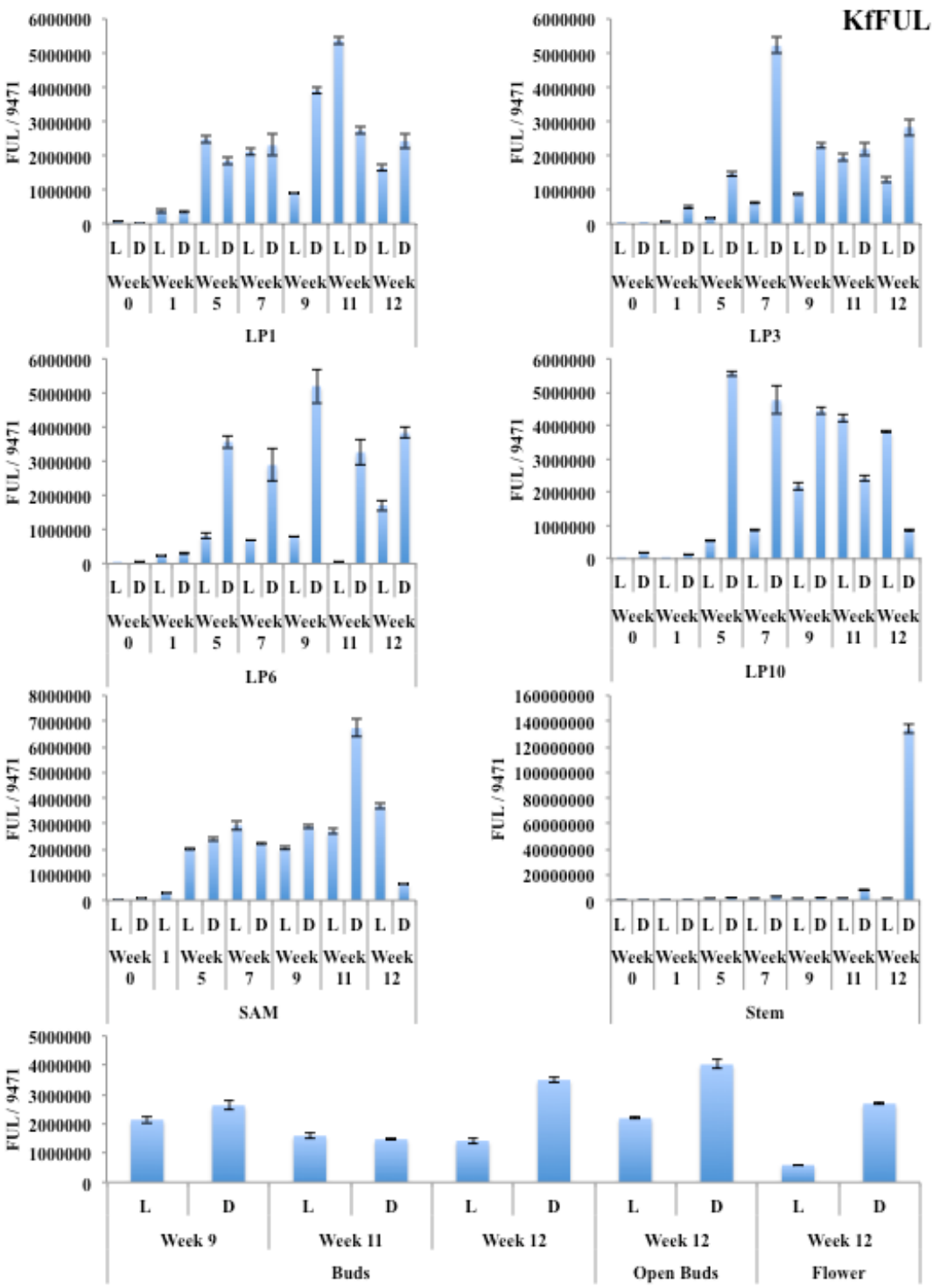

Figure 7.18. RT-qPCR data for *KfFUL* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

7.2.3 Changes of the transcript levels of flowering pathway genes in *KfCDF2* transgenic lines under non-flowering conditions

From observations using plants kept in the greenhouse under long-day conditions (16:8 LD), and also from the initial flowering time experiment (7.2.1), it had been demonstrated that the *KfCDF2_FL* over-expresser lines were able to initiate flowering, and also flower more rapidly, than wild type in all conditions. Also, in 12:12 LD conditions, as shown in Figures 7.5 and 7.6, the RNAi line developed apical hooks, whilst the wild type had yet to do so, suggesting these lines may also be able to develop flowers more rapidly under certain day-length conditions.

To investigate the regulation of genes hypothesised to function in the photoperiodic flowering time pathway, transgenic lines were grown for 6 months in 16:8 conditions in the greenhouse, and then placed in a large walk-in growth room under short-day conditions; where wild type plants had been shown previously to flower successfully. Wild Types used in this experiment also managed to produce flowers after 12 weeks. Interestingly though, these conditions prevented the *KfCDF2* transgenic lines from flowering, or even developing any signs that they were attempting to initiate flowering. The light spectrum between the greenhouse, where *KfCDF2_FL* lines flower in long days, the Snijders growth cabinets used for the initial experiment (7.2.1), and the walk-in growth chamber were therefore compared to assess the quality and quantity of light in each growth area (Fig 7.19).

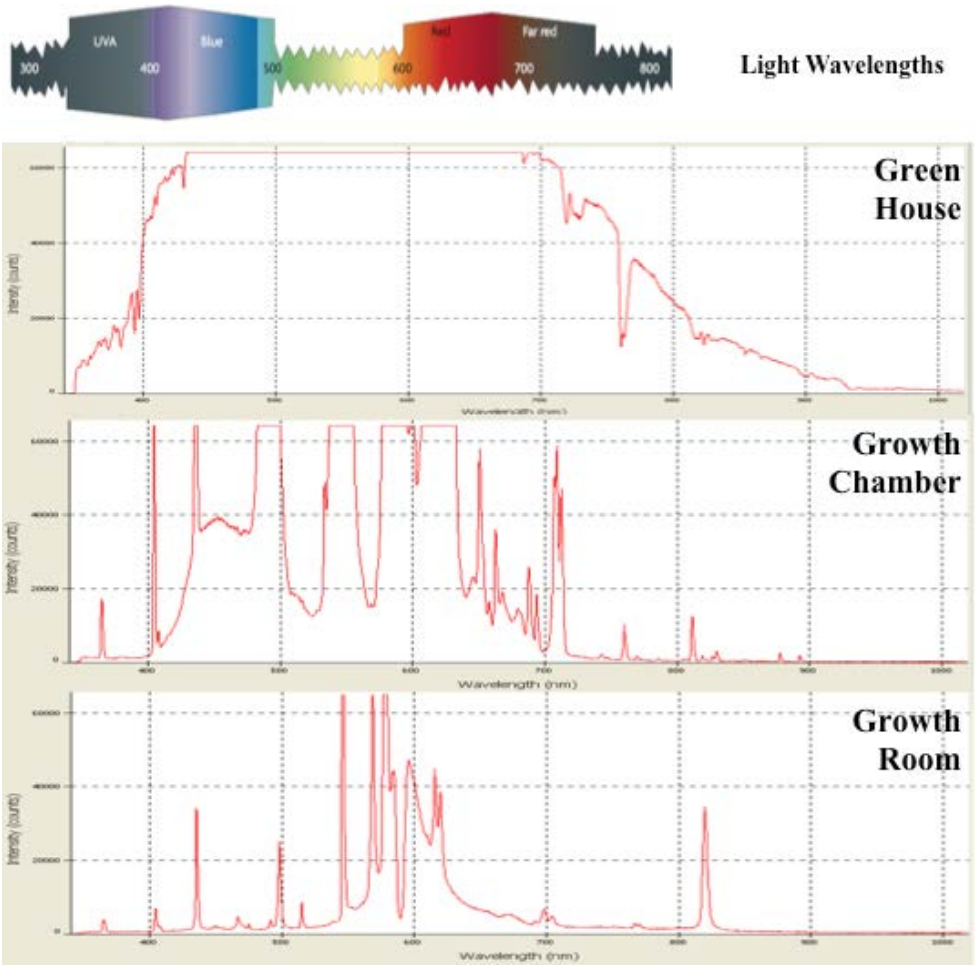

Figure 7.19. Light spectrum (quality of light) found in different growth areas. Greenhouse measurements were taken on an overcast day, during the morning and afternoon, but no difference in readings was detected. The morning reading is presented. Both greenhouse and growth chamber produced a better quality and quantity (intensity) of light than the growth room. Despite this, wild type plants still flowered in the growth room, suggesting that the *KfCDF2* transgenic lines were most susceptible to flowering inhibition when receiving either low quality and/ or quantity of light. Light Spectrum image taken from Jiao et al., (2007).

The light spectrum was compared within the three areas: greenhouse, Snijders Microclima MC-1000 growth chamber, and walk-in growth room containing high intensity white lights and also sodium lamps. As the growth room was the only area where *KfCDF2* transgenic lines did not flower, this spectrum was compared to that of the other two areas. The growth room had a lower quality of light compared to the other two areas (Fig. 7.19), as there were relatively low amounts of blue, red and far

red light emitted by the lamps in the growth room (Fig. 7.19; growth room). These wavelengths of light have been demonstrated in other species to be essential for the day-length dependent induction of flowering. It is thus noteworthy that wild type plants were able to flower under the growth room conditions. This suggests that changes in the transcript level of *KfCDF2* affected the plant's sensitivity to either blue and/ or red/far-red light. It therefore appears light quality is very important for *KfCDF2_FL* flowering. Furthermore, the quantity of light across most of the visible part of the spectrum was also lower in the growth room conditions, which could indicate that the *KfCDF2* transgenic lines required a high quantity of light in order to induce flowering.

Based on these findings, RT-qPCR was used to investigate the regulation of flowering pathway genes in the *KfCDF2* transgenic lines, *KfCDF2_FL_14C* and *KfCDF2_RNAi_19A* as well as the wild type in long days (16:8 LD) (Table 7.3 & Table. 7.4). All RT-qPCR graphs are presented in the appendix (S7.1 to S7.12) providing the full details of the magnitude of changes in transcript levels. To simplify the interrogation of this data, summary tables are presented here to highlight key trends within the gene transcript abundance data. In Tables 7.3 and 7.4, green squares represent up regulation compared to the wild type, red indicated down regulation and purple shown no significant change to the wild type.

7.2.3.1 Changes in the transcript abundance of flowering pathway genes in *KfCDF2_FL_14C*

In the *KfCDF2_FL_14C* line, *KfCDF2* was over-expressed in various tissues, with the only exceptions being LP3 and the dark period in the stem. This suggests that there

may be changes to *KfCDF2* regulation in different tissue types, depending on their importance in that specific area.

In *KfCDF2_FL_14C*, there were clearly genes that were up-regulated throughout the day and also throughout tissues, namely *KfCO25, KfCO46, KfAP1* and *KfFUL*, which are all genes that promote floral development, and must be up-regulated for flowering to occur. This could therefore explain why the *KfCDF2_FL* overexpressing plants flowered early, as they possessed elevated levels of these transcripts, which signalled that the external conditions are correct for flowering. As soon as these *KfCDF2_FL* plants were old enough, and thus had sufficient resources, it appears that flowering was initiated even in long day conditions. Other genes, such as *KfFT12, KfZTL* and *KfSOC1* also showed increased transcript levels at most times of day and in all tissues (Table 7.3). *KfFT18*, which from the wild type RT-qPCR data tracking the progression to flowering appeared to be involved in the repression of flowering, was also down-regulated in many leaf pairs/ organs, including LP1, LP3 and LP6, and the meristem in the light. This may explain why *KfFT12* and *KfAP1* were elevated in many of these leaf pairs/ organs, as levels of the putative repressor had been lowered. *KfSEP*, which is involved in controlling floral organ development in the different whorls of the developing floral meristem, was found to be under tight control. In the light period sampled, *KfSEP* transcript levels were elevated in all leaf pairs and organs, whereas in the dark sample this gene was down regulated. This could explain why flowering occurred more rapidly in these transgenic lines as a key gene involved in the regulation of the various whorls of floral organs was up-regulated, possibly leading to faster rates of floral organ development.

FD in Arabidopsis has been shown to interact with FT to induce many flowering genes, including *AP1*. It is therefore interesting that in this experiment in *K.*

fedtschenkoi, the *KfCDF2_FL* lines showed reductions in *KfFD* in half of the samples. However, based on understanding of the role and function of FD in Arabidopsis, FD is likely to interact with FT mainly in the stem and meristem, both of which only showed only very slight reductions compared to the wild type. Thus, these detected changes in *KfFD* levels may not impact on the progression to flowering in *K. fedtschenkoi*.

Two clock genes also had their expression profile analysed in this experiment, namely *KfGI* and *KfZTL*. In 10 out of the 12 tissue samples analysed, *KfZTL* transcript levels were up-regulated. In the other two samples (LP3 dark and LP10 light), there was no significant difference from the wild type. *KfGI* was only detected as being up-regulated in *KfCDF2_FL* in three samples, namely LP1 light and LP10 light and dark. *KfGI* transcript levels were down-regulated in the meristem in the dark, and in the stem in the light and dark, but showed no difference relative to the wild type in the remaining 6 samples. This could suggest that *KfCDF2* is able to directly or indirectly induce *CO*, as opposed to this induction requiring a mechanism involving *GI*.

Table 7.3. Summary of flowering pathway gene transcript abundance changes relative to wild type in *KfCDF2_FL_14C*.

CDF2	LP1		LP3		LP6		LP10		Meristem		Stem	
FL 14C	L	D	L	D	L	D	L	D	L	D	L	D
CDF2												
GI												
CO25												
CO46												
FT12												
FT18												
AP1												
SEP												
ZTL												
SOC1												
FD												
FUL												

Green indicates increased expression, whilst red shows a decrease. Purple indicates levels were not significantly different from those in the wild type.

7.2.3.2 Changes in the transcript abundance of flowering pathway genes in *KfCDF2_RNAi_19A*

With *KfCDF2_RNAi_19A*, the majority of tissues at the majority of times did show a reduction in the *KfCDF2* transcript level, consistent with the 35S promoter driving the hairpin RNA RNAi transgene in this line. The exceptions were LP1 in the light and dark, and the meristem and stem at the end of the light, which all showed an increased transcript level for *KfCDF2* compared to the wild type despite the presence of the *KfCDF2_RNAi* transgene (Table 7.4).

In this RNAi knockdown line, many of the other flowering pathway genes were again up-regulated in all samples, including *KfCO25, KfCO46, KfAP1* and *KfSOC1*, which are all involved in promoting flowering. *KfZTL* was also over-expressed relative to the wild type in every tissue and time-point, except for at the end of the light period in LP10. On the other hand, *KfGI* only showed over-expression in the light period for each of the sampled leaf pairs. *KfGI* also showed down-regulation in the stem and during the dark in the meristem. In the dark in LP1, LP3 and LP6 there was no difference to wild type for *KfGI*, whilst in LP10 there was a decrease in levels in the dark.

In contrast to the finding in *KfCDF2_FL_14C,* where *KfFUL* was consistently over-expressed, in the *RNAi_19A* line it showed no difference to wild type in any sample. This may therefore suggest that *KfCDF2* has a direct or indirect role in inducing *KfFUL* expression, with excess *KfCDF2* leading to higher *KfFUL* levels.

As in *KfCDF2_FL*, *KfSEP* was also up regulated in the light, but down regulated in the dark, consistently in all sampled tissues/ organs. *KfFT12* was only found to be up-regulated in the stem and meristem relative to wild type, whereas in most other tissues it was found to be at the same level as in the wild type, or even down-regulated in

LP1 and LP3 in the dark (Table 7.4). However, it must be remembered that increased transcript levels in the stem and meristem are likely to be the most important areas in terms of flowering, as this is where the mobile FT is required to induce flowering and may even induce its own expression in the stem and meristem as the FT protein becomes more abundant. *KfFT18* on the other hand was generally under-expressed or expressed at the same level in the various tissues/ organs of the RNAi line. Over-expression of the KfFT18 transcripts was only detected in LP1 at the end of the dark and LP10 in the light. *KfFD* expression in the meristem was down-regulated in the RNAi line, but again this was only a small change relative to the WT (Appendix S7.11). LP1 in the light and dark, LP3 and the stem in the dark, and LP6 in the light all showed increased transcript levels for *KfFD*, with the largest increase being detected in the stem.

Table. 7.4. Summary of changes in the transcript abundance levels of flowering pathway genes relative to wild type in the *KfCDF2_RNAi_19A* transgenic line

CDF2	LP1		LP3		LP6		LP10		Meristem		Stem	
RNAi 19A	L	D	L	D	L	D	L	D	L	D	L	D
CDF2												
GI												
CO25												
CO46												
FT12												
FT18												
AP1												
SEP												
ZTL												
SOC1												
FD												
FUL												

Green indicates increased expression, whilst red shows a decrease. Purple indicates levels do no differ from wild type.

As part of follow-up work aimed at determining the downstream target genes that *KfCDF2* binds to the promoters of, transgenic lines that over-expressed a HA-tagged

version of *KfCDF2_FL* were produced in *K. laxiflora* in the final 12 months of this project. *K. fedtschenkoi* and *K. laxiflora* are very closely related and very similar phenotypically, but *K. laxiflora* has been found to set viable seed when self-pollinated and so is believed to be a more amenable model species for further transgenic experiments. The two species have very similar genomes, although *K. laxiflora* is a tetraploid with a ~490 Mb genome whereas *K. fedtschenkoi* is a diploid with a 246 Mb genome.

Interstingly, the impact of overexpressing *HA-KfCDF2* in *K. laxiflora* appeared to be much stronger than when the same gene was over-expressed in *K. fedstchenkoi*. All of the *K. laxiflora* lines regenerated with the *HA-KfCDF2_FL* over-expression construct flowered early in long days in the greenhouse, indicating that the early flowering in long days phenotype displayed 100% penetrance for all screened and confirmed HA-*KFCDF2* overexpression lines in *K. laxiflora*. Samples have been collected from these lines, and hopefully the next step will be Chromatin Immunoprecipitation followed by high-throughput sequencing on the Illumina system (ChIP-Seq). This work should hopefully reveal exactly where this TF binds in the genome, and thus which downstream target genes are regulated by *KfCDF2*.

7.3 Discussion

As well as playing a role in the circadian optimisation of CAM (see Chapters 3 and 6), *KfCDF2* has also been identified here as playing a role in the photoperiodic regulation of the short-day-dependent flowering response in *K. fedtschenkoi* and *K. laxiflora*. *KfCDF2* over-expresser transgenic lines were found to flower constitutively in 16:8 LD greenhouse conditions (long days); conditions which have never been found to mediate flowering in wild type *K. fedtschenkoi*. The over-expresser lines were also shown to flower more rapidly than the wild type under both 12:12 and 8:16 light/dark conditions. They produced more flowers, and, visually at least, appeared to have fewer leaves, appearing to invest less energy into vegetative growth, and more into the reproductive effort.

Interestingly though, it appears that the quality and/or quantity of light affects the ability of the *KfCDF2* transgenic lines to flower. It was found that low levels of blue, red and far-red light in the visible spectrum and low light intensity, leading to flowering inhibition in the transgenic lines. This light quality and quantity effect did not affect the wild type though, which displayed a normal flowering response in 8:16 LD (short days). DOF transcription factors have been shown to play roles in phytochrome signalling, involving red/far-red light. An example of this is *COG1*, which when overexpressed leads to hyposensitivity to red and far-red light (Park *et al.,* 2003). This opens up the possibility that *KfCDF2* also plays a role in the phytochrome A and B signalling pathways. Phytochrome A mutants have also been shown to have delayed flowering in low irradiance (Bagnall *et al.,* 2001). Furthermore, because there was also less blue light in the growth room light spectrum, this may have led to ZTL and its subsequent complexes becoming less stable, which could lead to flowering inhibitors not being degraded. In relation to light quantity

during the day, the intensity was also found to be much lower in the growth room where the *KfCDF2* transgenic lines did not flower, compared to the green house and growth chambers where flowering was successful.

It is important to note in this context that SDPs have been found in general to be less sensitive to light quality and quantity (Park *et al.,* 2003). Instead, they are often characterised as being more sensitive to the length of the dark period, although this is not known specifically for the short day response in *K. fedtschenkoi*. Low irradiance levels will of course affect photosynthesis, and these plants have already been shown in the Chapter 6 to achieve reduced levels of dark CO_2 fixation relative to the wild type. This could therefore mean that they have fewer resources available to fuel the development of the floral meristem, suggesting that metabolic changes as a result of perturbations in photosynthesis could be the root cause of the inhibition of flowering in the *KfCDF2* transgenic lines under the growth room conditions used in the study.

There is a wealth of knowledge regarding the genetic and signal-transduction pathways involved in the regulation of flowering time in Arabidopsis, but much less is known about the equivalent pathways in short-day plants. Samples of all above-ground parts of the plants were therefore collected throughout the flower-induction process under short day conditions, and the transcript levels of flowering pathway genes were investigated in LP1, LP3, LP6, LP10, stem, and meristem, and later on buds and flowers. From this, gene transcript abundance could be monitored in the weeks leading up to the opening of the first flowers.

7.3.1 Flowering progression in wild type *K. fedtschenkoi*

The use of RT-qPCR to study the regulation of the steady-state transcript abundance of *KfCDF2* in wild type *K. fedtschenkoi* found that there was high transcript abundance for *KfCDF2* in long days in LP6 and LP10, but low transcript levels in other parts of the plant. This suggested that the increase in transcript levels in LP6 and LP10 may be due to CAM induction, where the *KfCDF2 TF* plays a role in the circadian control of CAM (Chapter 6). Once wild type plants entered short day conditions that induce flowering, the transcript abundance of *KfCDF2* was elevated in other tissues. From week 5, *KfCDF2* showed a large increase in the transcript abundance of *KfCDF2* in the stem and meristem, suggesting KfCDF2 may have different regulatory roles in different parts of the plant. In Arabidopsis, it has been demonstrated that *CDFs* repress *CO* expression in the leaves (Fornara *et al.,* 2009), but in the meristem *CDFs* may have a different role, possibly depending on which protein it is interacting with.

It is possible that *KfCDF2* functions as a promoter of flowering rather than a repressor in *K. fedtschenkoi*, as a sequence alignment of the amino acid sequences of the *KfCDF2* genes and the closest Arabidopsis ortholog revealed that they share only 38.8 % sequence identity. In addition, once the wild type plants had been transferred to short day conditions, the transcript abundance of *KfCDF2* dropped in older leaves, which could possibly be due to levels being reduced by GI and ZTL. This is supported by transcript data for both genes, showing increased transcript levels after a week of short day inductive conditions, which correlated with the reduction in *KfCDF2* transcript levels in the older leaf pairs. It would therefore be interesting to investigate the impact on CAM in *K. fedtschenkoi* that are progressing to flower in

short-days. If *KfCDF2* is down regulated, are other CAM genes also down regulated? Could flowering and reproduction be being prioritised over CAM?

The two *CO* sequences in *K. fedtschenkoi* shared 84.5% pairwise identity in an alignment of their nucleotide sequences, but showed peak transcript levels at different times of day. *KfCO25* tended to peak at the end of the light period, whereas *KfCO46* peaked during the dark. It therefore seems feasible that the ratio of expression between these two genes that leads to flowering could be the important factor for the induction of flowering in *K. fedtschenkoi*. *KfCO46* was also detected at high levels in buds, suggesting *KfCO46* could also function in the regulation of floral development, whereas *KfCO25* may perform more typical CO functions including mediating the induction of *FT* in short days.

For the two *FT* genes identified, the regulation of the transcript level of *KfFT18* was consistent with it acting as the antiflorigen, whilst *KfFT12* regulation was consistent with it functioning as the *K. fedtschenkoi* florigen; despite the homology to the Arabidopsis FT genes suggesting the opposite functions. *KfFT12* was found to increase in tissues the longer they were in short day 8:16 LD conditions, and was also found in developing buds, where *KfFT18* was barely expressed. The two genes also peaked at different times, which could suggest they regulate each other. *KfFT18* peaked during the light, suggesting that in long days lots of inhibitor is expressed preventing flowering. By contrast, *KfFT12* peaked during the dark, and so in longer nights more of the florigen would be able to be synthesised, leading to induction of flowering and could repress *KfFT18*.

KfAP1 was transcribed in all tissues even in non-flowering conditions, which is also the case in Arabidopsis (TAIR; www.arabidopsis.org), suggesting it does play other roles in the plant on a day-to-day basis beyond its best characterised role in the

regulation of floral development. However, *KfAP1* transcript levels increased by week 5/6 in all tissues. This is a developmental point at which the floral meristem had begun to develop, and, based on the known functions of its ortholog in Arabidopsis, *KfAP1* transcripts would be a key element in the development of the floral meristem (Teper-Bamnolker *et al.,* 2005). *KfAP1* also increased in transcript abundance in buds, suggesting this gene plays the same role in the regulation of flowering in *K. fedtschenkoi*, as it does for Arabidopsis, which is controlling sepal and petal differentiation and development (Mandel *et al.,* 1992; Kaufmann *et al.,* 2010).

KfSEP is also an important gene for flowering, and was repressed during the light period in non-inductive conditions. Under short-day inductive conditions, light period transcript levels for *KfSEP* were detected. *SEP2* and *SEP3* have been shown to function redundantly with *SEP1* (Pelaz et al., 2000), and are crucial for the transition to flower development in the meristem in Arabidopsis (Teper-Bamnolker *et al.,* 2005). Levels of *KfSEP* in *Kalanchoë* were found to be high in the meristem and also in buds, again suggesting this gene functions in a similar way to its Arabidopsis ortholog.

KfZTL appeared to play a role in the final floral development signal, as in all tissues at week 7, right before visible flowers were produced, there was a large increase in the transcript level of *KfZTL*. A high level of *KfZTL* transcript was detected in buds that were starting to open.

Despite *KfSOC1* being a flowering gene, it has also been shown recently to play a role in stomatal control (Kimura *et al.,* 2015). This could explain why in leaf pairs 6 and 10, levels were high under non-inductive long-day conditions. Once plants were transferred to short-days and the progression to flowering had commenced, the levels of this transcript dropped. This could again suggest that flowering may be prioritised

over CAM. It would be of interest to examine stomatal behaviour in leaves during the short-day induction of flowering in *K. fedtschenkoi*.

KfFD transcript regulation did not correlate well with FD's proposed function in Arabidopsis. FD has been shown to interact with FT to induce flowering. *KfFT12* was not really detected until at least week 5 in all tissues, whereas FD expression greatly drops after week 5. It may indicate that only small levels of the two proteins are needed to induce flowering, or could suggest a different pathway is used. Another suggestion could also be that as *KfFT18* has the highest homology to the *Arabidopsis* florigen, then *KfFD* may in fact interact with that instead, and so may have taken on a flowering repression role.

KfFUL transcript regulation displayed a similar pattern of transcript levels to its Arabidopsis counterpart. It was barely detected in non-inductive long day conditions, but increased in all tissues under short days from week 5, and also increased as the buds developed, supporting its proposed role as a gene that promotes flowering. Interestingly, in *Arabidopsis*, *LFY* is a very important gene for flowering, as it was shown to be both necessary and sufficient to induce both flowers and leaves (Blazquez *et al.,* 1997). A *KfLFY* gene candidate (KF07545) was found in the *K. fedtschenkoi* genome, and was amplified successfully from genomic DNA, but it was not detected as an expressed transcript in both an existing database of RNA-seq quantitative transcriptomic data, and also after 40 cycles of RT-qPCR. This could suggest that an alternative gene performs the role of *LFY* in *K. fedtschenkoi*, or it may be that it can achieve its required function in *K. fedtschenkoi* from a very low transcript level that could not be detected with the methods used. Unfortunately, the role of *KfLFY* in the *K. fedtschenkoi* flowering pathway could not be determined, and

so this may suggest that *KfAP1* is able to function alone to initiate the development of flowers in this species.

7.3.2 The impact of changes in *KfCDF2* expression on the regulation of genes associated with the flowering pathway

With the regulation of flowering pathway genes in wild type in response to short days established as a baseline, it was then of interest to investigate the same genes for possible differential regulation in the *KfCDF2_FL* line that was able to initiate flowering in good quality greenhouse light under long days, without receiving a short day flowering signal. As *KfCDF2_RNAi_19A* also showed a more rapid stem hooking phenotype in 12:12 LD conditions relative to wild type, it seemed that this down-regulated *KfCDF2* transgenic lines was also able to flower more rapidly in response to short days, but only if given the initial flowering stimulus, which the *KfCDF2_FL* lines no longer require.

In both the *KfCDF2_FL* and *RNAi* line, many flowering pathway genes were up-regulated constantly; in particular *KfCO25, KfCO46* and *KfAP1*. This data provides a molecular-genetic basis to explain why the *KfCDF2* over-expresser line could flower more rapidly, and also suggests why, under 12:12 conditions, the *KfCDF2_RNAi* line was also primed to flower. For *KfCDF2_FL_14C*, there were many more genes up regulated in many more tissues than in the *RNAi_19A* line. Constitutive overexpression of *KfFUL* was also detected in the FL line, and constantly high transcript levels of *KfSEP* were detected in the light. The accumulation of *KfFUL*, *KfSEP3* and *KfAP1* in the meristem is the key step in induction of the floral meristem in *K. fedtschenkoi*. *KfCDF2_FL* had a constantly high level of these transcripts, which suggests that the *KfCDF2_FL* over-expresser line is constantly receiving a flower-

inducing signal, and that plants are simply waiting until they have accumulated sufficient storage carbohydrate/ lipid and/ or protein reserves to be able to initiate flowering.

By contrast, in the *KfCDF2_RNAi_19A*, transcript abundance for *KfFUL* was similar to the wild type, which is consistent with the fact that these lines did not initiate flowering in 16:8 LD conditions (long days). In the *KfCDF2_RNAi* line, *KfCDF2* expression was found to increase in the meristem and stem in the light, but this was not sufficient to signal an increase in *KfFUL*, even though there was an increase in *KfFT12*. This suggests that a threshold level of *KfCDF2* is required in order to induce *KfFUL* expression, which would in turn initiate the transition to flowering. This could also help to explain why some of the independent *KfCDF2_FL* over-expresser lines only progressed as far as producing hooked stems under 16:8 greenhouse conditions, but did not actually succeed in flowering. This could be due to these lines having a level of *KfCDF2* over-expression that was not quite sufficient to cause an increase in *KfFUL* expression to the required level.

In the *KfCDF2_RNAi* lines, *KfSOC1* was over-expressed constantly in all tissues sampled. This could be a further reason why these lines are not flowering independently of the photoperiod, because in non-inductive long day conditions, *KfSOC1* was over-expressed, which may have prevented the progression to inflorescence meristem initiation (Lee *et al.,* 2010).

Finally, it was surprising that both *KfCDF2* transgenic lines lacked a detectable level of transcripts for the *KfLFY* gene that was identified in the genome. This suggests that *KfLFY* was not expressed in *K. fedtschenkoi* under the conditions studied here using the RT-qPCR technique. It would appear from these findings that *KfLFY* does not play a role in the regulation of floral development in response to short days in *K.*

fedtschenkoi. This is certainly an area that merits further work. A more comprehensive RT-qPCR or quantitative RNA-seq study should be undertaken to determine whether or not the *KfLFY* gene is expressed under any conditions and / or in any specific tissue / organ during the development of a *K. fedtschenkoi* plant. It is noteworthy in this context that an assembled *K. laxiflora LEAFY* transcript was detected in a PASA (Program to Assemble Spliced Alignments) assembly of *K. laxiflora* transcripts that were assembled using Illumina Hi-Seq reads from a wide range of tissues, organs and developmental stages (Won Cheol Lim, Hengfu Yin, Xiaohan Yang, John Cushman and James Hartwell, unpublished findings). This suggests that *KlLFY* was being expressed in some of the wide range of *K. laxiflora* tissues and developmental time points that were sampled for the Illumina sequencing that led to the *K. laxiflora* PASA transcriptome assembly. These findings further emphasise the importance of future work to identify which tissues/ organs and/ or developmental stages possess detectable levels of *KlLFY* or *KfLFY* expression in *K. laxiflora* and *K. fedtschenkoi* respectively.

7.4 Conclusions

There were many similarities between the patterns of gene regulation observed during the transition to flowering in short days for *K. fedtschenkoi* studied here and previous studies with Arabidopsis. This is interesting considering that Arabidopsis is a LDP and *K. fedtschenkoi* is a SDP. The transcript abundance profiles measured here indicate that many flowering pathway genes are likely to carry out similar functions in both species, including *KfAP1, KfCO* and *KfFUL*. Other flowering pathway genes appeared to function in a different way in *K. fedtschenkoi* relative to the known function and pattern of regulation in Arabidopsis, including *KfFD* and *KfLFY*; with the latter gene not being detected as expressed in any of the sampled tissues or time-points. *KfCDF2* was one of the genes that was found to conduct a different function, as in LDPs *CDF2* overexpression has been shown to delay flowering time, rather than initiating and accelerating it (Fornara *et al.,* 2009).

Much more work needs to be conducted in this area to establish the exact functions of many of these flowering genes, and also establish changes associated with day length and also CAM, which *KfCDF2* has been demonstrated to play a role in regulating (Chapter 6). Furthermore, the findings reported in this chapter, and the recently published findings of Kimura *et al.* (2015), suggest that *KfSOC1* could be involved in stomatal regulation in CAM performing leaves of *K. fedtschenkoi* where it was up-regulated, and this TF also merits detailed further investigation through the development of transgenic over-expresser and RNAi lines, and ChIP-Seq analysis. The *KfCDF2* transgenic lines were found to be more sensitive to light quality or quantity than wild type lines, as they did not flowering in growth room short day conditions under which the wild type did successfully flower. It is interesting in this context that some *DOF* factors have already been shown to be involved in light

sensing and signalling pathways in Arabidopsis (Park *et al.,* 2003). Further work needs to be carried out, with careful control of both the quantity and quality of the light, in order to determine whether light quality or quantity is important for the *KfCDF2* transgenic lines to flower early. If light quality is most important then experiments using different LED lamps that emit very specific wavelengths of light should permit a precise dissection of the optimal wavelength(s) of light required for the induction of flowering through a *KfCDF2* signalling pathway in response to short days.

ChIP-Seq analysis of the genome wide promoter and other gene regulatory regions bound by *KfCDF2* will also be a critical step towards understanding which genes are bound to and regulated by *KfCDF2*. These results will shed more light on the direct functions of *KfCDF2*. It will also be interesting to investigate the upstream promoter region of the *KfCDF2* gene itself, in order to identify the important promoter motifs required for its induction in leaf pair 6 relative to leaf pair 1, and its light/ dark and circadian clock control. This could be pursued using promoter deletion analysis with a series of *KfCDF2* promoter::luciferase reporter constructs. Luciferase imaging of stable transgenic lines carrying different length fragments of the *KfCDF2* promoter should allow the minimal promoter for leaf pair 6 induction, and light/ dark and circadian control to be identified.

It was also interesting to discover during the last few months of the practical work on this project that the newly generated *HA-KfCDF2* over-expresser lines in *K. laxiflora* all progressed to flowering in 16:8 long day conditions in the greenhouse. This suggested that *K. laxiflora* may be more susceptible to over-expression of *KfCDF2* than *K. fedtschenkoi*, as only the strongest *KfCDF2* over-expresser lines in *K. fedtschenkoi* flowered constitutively in long day greenhouse conditions. However, it

is possible that all of the more recently generated *K. laxiflora* lines over-express *KfCDF2* more strongly that most of the previous population of *K. fedtschenkoi KfCDF2* over-expresser lines and further work must now be undertaken to characterise the level of transgene expression in the new *HA-KfCDF2* over-expresser lines of *K. laxiflora*. However, it remains possible that flowering in *K. laxiflora* was more susceptible to elevated levels of *KfCDF2*. It is thought provoking to speculate that this may be in some way be linked to the fact that *K. laxiflora* is able to develop viable seed following self pollination, whereas *K. fedtschenkoi* fails to form viable seed due to a mutation in the *LEAFY-COTYLEDON1* (*LEC1*) gene (Garces *et al.,* 2007). The *LEC1* mutation found in several plantlet-forming species of *Kalanchoë* causes a C-terminal truncation of the *LEC1* protein, and results in a failure of developing seed to achieve dehydration tolerance in the final stages of seed maturation. Whether or not this mutation also impacts on the flowering pathway and the role of *KfCDF2* in that signalling pathway is an interesting area for future investigation, which is readily tractable via a direct comparison of flowering pathway signalling in *K. fedtschenkoi*/ *K. laxiflora,* as the two species possess a mutated *LEC1* and a functional *LEC1,* respectively.

(Hernando-Amado *et al.,* 2012). The first DOF was identified in 1993 by Yanagisawa *et al*., and since that first report many more plant *DOF*s have seen identified, characterised and analysed in terms of their sequence homology (Yanagisawa 1995; Vicente-Carbajosa *et al.,* 1997; Plesch *et al.,* 2001), protein interactions (Zhang *et al.,* 1995; Fornara *et al.,* 2015), binding to specific DNA elements (Kisu *et al.,* 1998; Baumann *et al.,* 1999; Fornara *et al.,* 2009), *in planta* functions (Negi *et al.,* 2013; Corrales *et al.,* 2014) and abnormal phenotypes and patterns of expression (Papi *et al.,* 2002; Park *et al.,* 2003). From these *in vivo* and *in vitro* experiments, a variety of putative roles have been suggested for the *DOF* family of TFs in plants (Table 6.1). In Arabidopsis, the vast majority of the biological functions that have been identified as being controlled by this TF family have been attributed to large-scale genome duplication events that have enabled these transcription factors to adapt and evolve new roles (Vision *et al.,* 2000b). Phylogenetic trees generated based on protein and nucleotide sequence alignments for all known *DOF* genes have suggested that this gene family multiplied prior to angiosperm diversification (Vision *et al.,* 2000a). Multiple *DOF*s have also been found in gymnosperms, which are thought to have diverged from the angiosperms around 245-202 million years ago (Hochuli *et al.,* 2013). The *DOF* family have also been found in lower plants like mosses (*Physcomitrella*) and green algae, although the green alga *Chlamydomonas reinhardtii* only possesses one *DOF* TF, which has therefore been proposed as representing the original *DOF* TF. Interestingly, the genomes of other lower photosynthetic eukaryotes, including diatoms and red algae, have not been found to possess this *DOF* family of TFs, suggesting they are have only been present since the evolution of the higher green photosynthetic eukaryotes (Yanagisawa 2002; Shigyo *et al.,* 2007).

Table 6.1 Examples of the putative biological functions for members of the DOF transcription factor family in plants, showing a large and diverse range of functions.

DOF putative function	Reference(s)
Light response	(Yanagisawa *et al.,* 1998; Santopolo *et al.,* 2015)
Defence response	(Zhang *et al.,* 1995; Sasaki *et al.,* 2015)
Seed germination	(Papi *et al.,* 2000; Gualberti *et al.,* 2002)
Auxin response	(Kisu *et al.,* 1998; Baumann *et al.,* 1999; Sagar *et al.,* 2013)
Gibberellin response	(Washio 2001)
Guard cell specificity	(Plesch *et al.,* 2001; Negi *et al.,* 2013)
Flowering regulation	(Imaizumi *et al.,* 2005; Fornara *et al.,* 2009; Corrales *et al.,* 2014)
Carbon metabolism	(Yanagisawa *et al.,* 1998; Umemura *et al.,* 2004)
Abiotic stress response	(Park *et al.,* 2003; Corrales *et al.,* 2014)

From intensive research over the last two decades, *DOF* transcription factors have now been demonstrated to function either to induce and/ or repress target genes through binding to their upstream promoter regions, and have also been demonstrated to have redundant functions, whereby several closely related *DOF*s all function redundantly in a single biological pathway (Fornara *et al.,* 2009; Fornara *et al.,* 2015). Interestingly, other experiments have shown that certain closely related *DOF* TFs can also carry out opposing roles; for example, *DAG1* and *DAG2* in Arabidopsis. When *DAG1* was disrupted it resulted in the loss of seed dormancy, as the seeds of the *dag1* mutant did not require light for germination (Papi *et al.,* 2000). By contrast, the *dag2* mutant had the opposite effect, as the seeds were more dependent on a light stimulus for the initiation of germination (Gualberti *et al.,* 2002).

6.1.1.1. Previously known DOF roles suggesting TFs may play a role in the circadian optimisation of CAM

As well as the results from the preliminary *K. fedtschenkoi* SOLiD RNA-seq data for C_3 and CAM leaves sampled over the light/ dark cycle, and semi-quantitative RT-

PCR data for constant light (LL) conditions presented previously in Chapter 3, there are also a wide range of published reports describing the regulation and functions of *DOF* transcription factors. Many of the published articles also reveal that this family can play an important role in the circadian optimisation of plant biology. In maize, a species that uses the C_4 adaptation of photosynthesis, *Dof1* has been shown to bind to the promoter of the C_4 specific isogene of PEPC, and control its light/ dark pattern of expression (Yanagisawa *et al.,* 1998; Yanagisawa 2000). Similar findings have also been made for following the constituitive over-expression of maize *Dof1* in Arabidopsis (Yanagisawa et al., 2004). Transgenic expression of maize *Dof1* in Arabidopsis resulted in the up-regulation of *PEPC* and pyruvate kinase genes, along with associated changes to central carbon and nitrogen metabolism, including a marked and significant reduction in the level of free glucose (Yanagisawa *et al.,* 2004). Another member of the Arabidopsis *DOF* family, called *STOMATAL CARPENTER1* (*SCAP1*), has also been found to be essential for the normal development of stomata (Negi *et al*., 2013). Mutant *scap1* lines showed irregularly shaped guard cells, and also displayed a failure of light-induced stomatal opening and CO_2 induced closing (Negi *et al.,* 2013).

6.1.1.2 DOF protein structure

The proteins of the DOF family are typically between 200 and 400 amino acids, and share a high level of sequence homology at their N-terminal end (the DOF DNA-binding domain), whilst the rest of the amino acids are divergent (Yanagisawa 2002). All DOF transcription factors share the DNA-binding amino acid sequence: $CX_2CX_{21}CX_2C$ (Fig 6.1A). This high homology demonstrated in the DNA-binding domain suggests that DOF proteins all have a similar DNA-binding specificity, and,

unlike many other families of transcription factor, including MYBs, the DOF family only possesses a single copy of the DOF domain. From the highly conserved amino acid sequence, a secondary structure was proposed, suggesting that these TFs possess a single zinc finger, which can been seen in Figure 6.1B (Yanagisawa 1995; Yanagisawa *et al.,* 1998).

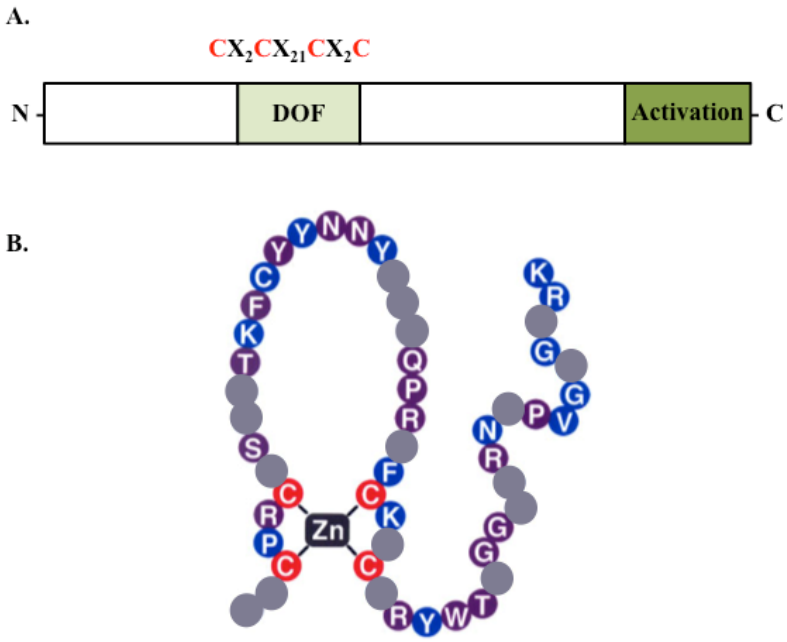

Figure 6.1. Example of a DOF proteins primary and secondary structure. **A**. shows primary structure with the DOF domain (light green) close to the N-terminus of the protein, and the activation domain (dark green), at the C-terminus of the protein. Close to the DOF domain should be the nuclear localisation signal, and also a serine rich region. Above the DOF domain is the key 29aa sequence present in all DOF TFs, in which the 4 C's interact to form the Zinc finger domain (X represents any amino acid). **B.** Displaying the proposed secondary structure of the 52aa DOF domain. Image has been edited from Yanagisawa (2004). This shows the four key Cysteine (C) residues (red) and other 100% conserved amino acids (purple) and some well conserved residues (with at least 75% conservation) are shown in blue, that together produce the one Zinc (Zn) finger DNA-binding domain. Grey circles indicate redundant residues.

Downstream of the DNA-binding domain, it is common to find serine residues, which are thought to be the 'molecular hinges' linking the DOF and activation domains (Yanagisawa 2002). The DOF family of proteins have been shown to conduct 'dual

functions', with a bi-functional domain for DNA-binding and also protein-protein interactions (Yanagisawa 1997). The first DOF interactions were reported to occur with a bZIP TF by Zhang *et al.*, in Arabidopsis (1995). The DOF protein OBP1 was shown to interact with bZIPs OBF4 and 5, and enhanced bZIP binding through the DOF domain *in vitro* (Chen et al., 1996). When promoter sequences were investigated *in vivo*, binding sites for OBP1 and OBFs were found to be in close proximity to one another. Since then, other interactions have also been found with other bZIPs (Albani *et al.,* 1997; Vicente-Carbajosa *et al.,* 1997), as well as MYBs (Diaz *et al.,* 2002) and WRKY transcription factors (Zou *et al.,* 2008) to name just a few.

6.1.2 Functional diversity of DOFs

As the DOF domain is conserved between the transcription factors in the DOF family, it has been demonstrated that the family show affinities for the same DNA-binding sites (Plesch *et al.,* 2001). Research has shown that AAAG or its reverse sequence CTTT are found in the target binding sequences for most individual DOF proteins (Yanagisawa *et al.,* 1993; Plesch *et al.,* 2001). In maize, it was shown that different DOF TFs achieved differential binding to distinct target promoters, according to the sequence flanking the core motif (Yanagisawa 1997). Another mechanism that allows for differential binding is protein-protein interactions, which occur through the activation domain. The interaction with individual DOFs with other transcription factors allows their binding targets and activation state to be modulated (Yanagisawa 1997).

A further protein interaction of DOFs has been reported to occur through interaction with high mobility group (HMG) proteins. HMG proteins are abundant, non-histone nuclear proteins, which function in DNA-dependant processes, such as transcription

and DNA replication and repair (Bustin *et al.,* 1990; Bustin 1999). These proteins have been shown to interact between the DOF domain and HMG box. In maize, Dof1 and 2 have been shown to have their DNA binding activity enhanced when they are attached to certain HMGs (Yanagisawa 1997). It was also demonstrated that different HMGs interact with different DOFs, and also various HMGs promote binding to different extents (Krohn *et al.,* 2002). Furthermore, the phosphorylation state of the HMG was found to be important; as when the HMG was phosphorylated there was a weaker enhancement of DNA-binding (Bustin *et al.,* 1990).

6.1.3. *KfCDF2*

In *K. fedtschenkoi, KfCDF2* overexpressing lines, as mentioned in Chapter 3, showed photoperiod-independent flowering, with higher overexpressing lines producing flowers and lower overexpression lines only producing hooked meristems. This therefore suggested that *KfCDF2* plays a role in flowering time regulation in response to changing day-length, which is in-turn coupled to the central circadian clock. The role of *KfCDF2* in the regulation of flowering will be discussed in more detail in the next chapter (Chapter 7). This chapter will focus on the impact of *KfCDF2* on CAM and its regulation over the light/ dark cycle and under circadian free-running conditions.

6.2 Results

6.2.1 Bioinformatic characterisation of *KfCDF2*

To try and determine whether *KfCDF2* may play similar roles to Arabidopsis *CDF2* (At5g39660) the two sequences were translated, and the full-length amino acid sequences were aligned using the default Geneious Pro 4.5.5 global alignment tool; Blosum62 (Fig. 6.2). As reported previously for many other DOFs in a range of plant species, it was found that there was very little conservation outside the DOF and activation domains (Fig. 6.2). The amino acid sequences of KfCDF2 and AtCDF2 were found to share only 38.8% pairwise identity. Within the DOF domain though, the two proteins shared a high level of sequence identity (86.9%), with both DOF domains maintaining the 4 critical cysteine residues for the formation of the zinc finger, in addition to a number of other conserved residues (Fig. 6.1B), which play an important role in the structure of the DOF domain (Fig. 6.2B).

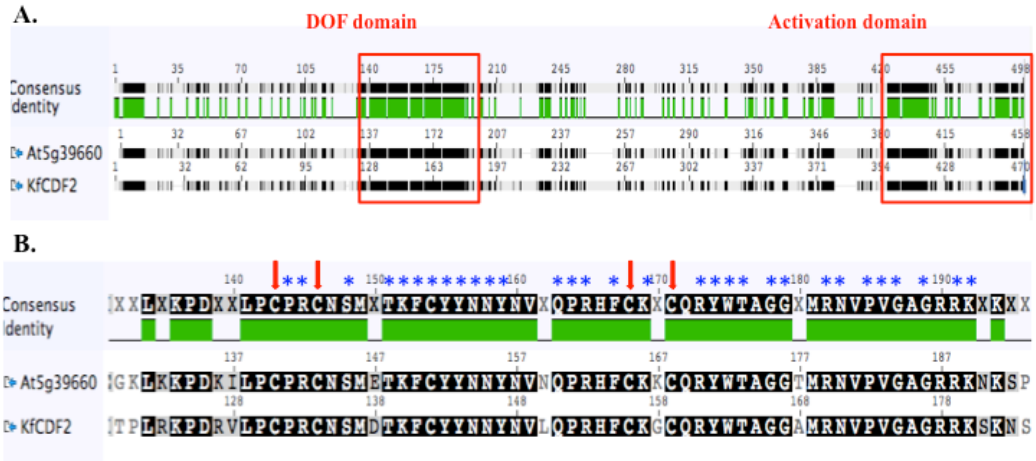

Figure 6.2. Global amino acid sequence alignment between KfCDF2 and its Arabidopsis ortholog AtCDF2 (AT5G39660) **A**. KfCDF2 and AtCDF2 both show a high level of sequence conservation within their DOF and activation domain, as highlighted by the red boxes. **B**. A close-up view of the amino acid alignment withint the DOF domain. Red arrows indicate the 4 conserved cysteine residues required for the formation of the DNA-binding zinc finger. Blue stars show all of the highly conserved residues shared by all plant DOFs.

As DOF family TFs have been shown to act redundantly in a number of other species, a second *KfCDF-related* gene was identified from the assembled genome and transcriptome databases. This second gene matched *AtCDF3* in Arabidopsis (At3g47500). A sequence alignment between the amino acid sequences of the two *KfCDF* genes revealed that they shared 76.4% pairwise identity, with large numbers of conserved amino acid residues outside of the DOF and activation domains. The level of divergence between *KfCDF2* and *KfCDF3* suggests that in *K. fedtschenkoi*, redundancy of *CDF*-related TFs is also likely to occur to some extent.

As mentioned in previous chapters, a large Illumina RNA-seq dataset was available defining gene regulation in CAM leaves of several previously characterised transgenic lines of *K. fedtschenkoi,* which had been manipulated for genes known to play a role in either the CAM pathway (RNAi lines targeting *KfPPCK1*, *KfNAD_ME_a1*, or *KfPPDK*), or the central circadian clock (*McTOC1* over-expressor line) (Boxall, Dever and Hartwell, unpublished data). This dataset enabled the impact of changes in the transcript abundance of known CAM and clock genes on the regulation of *KfCDF2* over the light/dark cycle in CAM leaves to be investigated in more detail.

The Illumina RNA-seq data for *KfCDF2* revealed that the FPKM values varied little during the light period, which was consistent with the low level of *KfCDF2* transcripts throughout the light period in the wild type (Fig. 6.3). In contrast, during the dark period, especially when *KfCDF2* rises progressively to its daily peak through 18:00, peaking at 22:00 (2 h before dawn) there were large differences in the transcript abundance of *KfCDF2* between the wild type and all of the transgenic lines (Fig. 6.3). In the wild type, *KfCDF2* shows the highest FPKM value at 22:00, when

maximum expression occurs in all lines, with an average FPKM value just below 800 (759 FPKM; Fig. 6.3). All other lines showed reductions in their FPKM values relative to the wild type at 22:00 (Fig. 6.3). Amongst the transgenic lines, *KfCDF2* had the highest transcript abundance level at 22:00 in *rPPDK1*, where the FPKM was only reduced to just under 600 (583 FPKM; Fig. 6.3). All other lines: *TOC1_OX* (197 FPKM), *rNAD-ME* (273 FPKM) and *rPPCK1_3* (220 FPKM) had FPKM values that were between 3.9-fold (*TOC1-OX*) and 2.8-fold (*rNAD-ME1*) lower than the wild type value at 22:00 (Fig. 6.3).

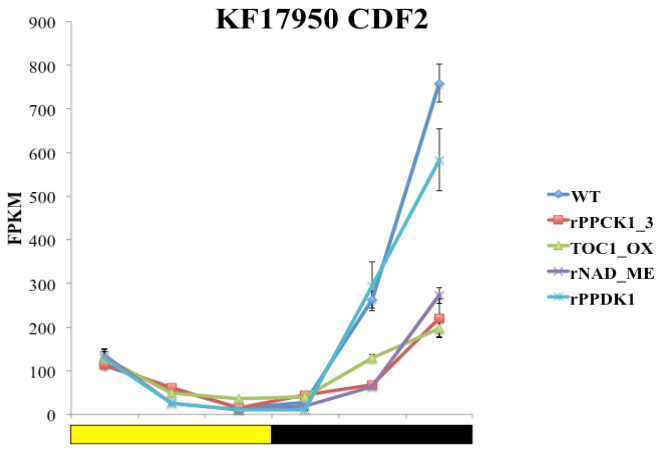

Figure 6.3. Differential regulation of *KfCDF2* in wild type and transgenic lines of *K. fedtschenkoi*. Quantitative RNA-seq data (FPKM values) is presented showing the light/ dark oscillation in the transcript abundance of *KfCDF2* in wild type and transgenic lines. Transgenic lines included the RNAi silenced lines for the three key CAM genes: *rPPCK1_3, rNAD-ME* and *rPPDK*, plus an over-expressor line that expresses the central clock gene *McTOC1* constitutively. *KfCDF2* transcript levels were reduced at the time of peak transcript (22:00 dark, 2 h before dawn) in all transgenic lines, but were reduced up to 4-fold in lines *rPPCK1_3*, *rNAD_ME1* and *TOC1_OX,* which all fail to maintain robust oscillations in circadian clock-controlled genes in LL constant light conditions.

The large reduction in the transcript abundance of *KfCDF2* in these transgenic lines supports the original hypothesis that *KfCDF2* is involved in an output pathway from the clock. Arrhythmia, and/ or dampening of oscillations of gene expression have often been observed in other plant species when core clock genes such as *TOC1* were

over-expressed (Dall'omo 2011). Furthermore, the fact that *KfCDF2* transcript levels were drastically lower in the *rNAD-ME1* and *rPPCK1_3* lines suggests that those genes may be involved in some form of feedback regulation, either directly, or indirectly via changes in metabolism associated with the perturbations in CAM-associated metabolism/ metabolites in these lines, that signals to regulate the level of *KfCDF2* transcripts in leaf pair 6. In particular, the existing phenotypic characterisation of these lines has demonstrated that whilst *rPPDK* is able to maintain robust circadian oscillations of clock-controlled genes under LL constant light conditions, both *rNAD_ME1* and *rPPCK1_3* displayed arrhythmia for several clock and clock-controlled CAM-associated genes when placed under LL free-running conditions (Dever *et al*., 2015; Boxall *et al*., 2015, manuscript in preparation). The low transcript abundance of *KfCDF2* at 22:00 in 12:12 LD cycles shown in Figure 6.3, may either be a cause or consequence of a lack of ability of these lines to maintain robust rhythms under LL conditions.

6.2.2 The impact of transgenic perturbation of *KfCDF2* levels on the light/dark regulation of key CAM- and circadian clock-associated genes

Semi-quantitative RT-PCR analysis was conducted to determine the effects of mis-expressing *KfCDF2* on other core CAM-associated and circadian clock-associated genes when compared to wild type *K. fedtschenkoi*. Leaf pair 6 (LP6; full CAM leaves in wild type) were collected every 4 h over a 12:12 light/dark cycle. Genes screened included the core clock genes *KfCCA1* and *KfTOC1*, the CAM carboxylation pathway genes *KfGWD, KfMEX1, KfGPT2, KfPPCK1* and *KfPEPC*, and the CAM decarboxylation pathway genes *KfPPDK, KfPPDK-RP, KfNAD-ME α,* and *KfNAD-ME β.*

6.2.2.1 Light/dark regulation of CAM and Clock genes in *KfCDF2* full-length over-expression lines

Despite transgenic expression of the full length open reading frame of *KfCDF2* being driven by the CaMV35S promoter, the transgene was still found to oscillate in transcript abundance over the 24 h LD cycle (Figs. 6.4 & 6.5); supporting the results already reported for this line in Chapter 3.2.4. Both *KfCDF2* over-expresser lines showed light/ dark transcript abundance patterns similar to the wild type, except with a higher level of *KfCDF2* transcripts at each time point (Figs. 6.4 & 6.5). *KfCDF2_FL_13A* over-expressed *KfCDF2* much more strongly than *KfCDF2_FL_14C*; consistent with the initial *KfCDF2_FL* screens (Chapter 3.2.4). In particular, line *KfCDF2_FL_13A* displayed very high levels of *KfCDF2* throughout the dark period. Specifically, line *FL_13A* reached transcript levels 28.8-fold higher than the wild type at 14:00 (2 h into the dark period), at a time when the wild type has yet to begin its nocturnal rise in *KfCDF2* transcript levels.

This more detailed time course experiment allowed more accurate determination of the fold-increases in *KfCDF2* transcript levels throughout the LD cycle. In contrast to the initial screening results reported in Chapter 3, *KfCDF2_FL_13A* showed on average a 3.7-fold increase in the light, and a 5.6-fold increase in the dark, whereas *KfCDF2_FL_14C* displayed a 2.1-fold average increase in the light, and a 1.4-fold average increase during the dark. On average, there was a larger amount of *KfCDF2* transcript produced in *KfCDF2_FL_13A*, and so, assuming there is a dosage response effect of changes in the level of this gene and its encoded protein, it was predicted that there would be greater knock-on effects on other processes in this line relative to line *14C*. Whereas *13A* showed only very small changes in the transcript abundance and temporal regulation for genes measured in this study (Fig. 6.4), line *14C* showed

several larger changes. Relative to the wild type, the biggest differences seen were in CAM-associated genes in both the carboxylation and decarboxylation pathways. The sugar and starch metabolism genes also showed small changes in abundance.

For both over-expressing lines, the transcript abundance of *KfPPCK1* was lower at its peak in the middle of the dark period, especially for line *14C* (Fig. 6.5). For most other genes, line *13A* had a small increase in expression, whereas *14C* displayed a decrease in the transcript abundance of the same genes (Fig. 6.4). For line *13A*, at 14:00 (2 h into the 12 h dark period) when *KfCDF2* had not yet begun to increase in the wild type, but was high in *FL_13A*, there was a small peak visible in many of the CAM-associated genes. For all of the genes, which show increases in expression in 13A at 18:00, all show a drop in expression level in *FL_14C*. These results therefore suggest that *KfCDF2* could be having some sort of indirect effect on these CAM-associated genes. As the light/ dark oscillations of the central clock genes *KfCCA1* and *KfTOC1* were perturbed very little in either *KfCDF2_FL* line, it is unlikely that *KfCDF2* levels feed back directly to impact on regulation of genes within the core

oscillator mechanism.

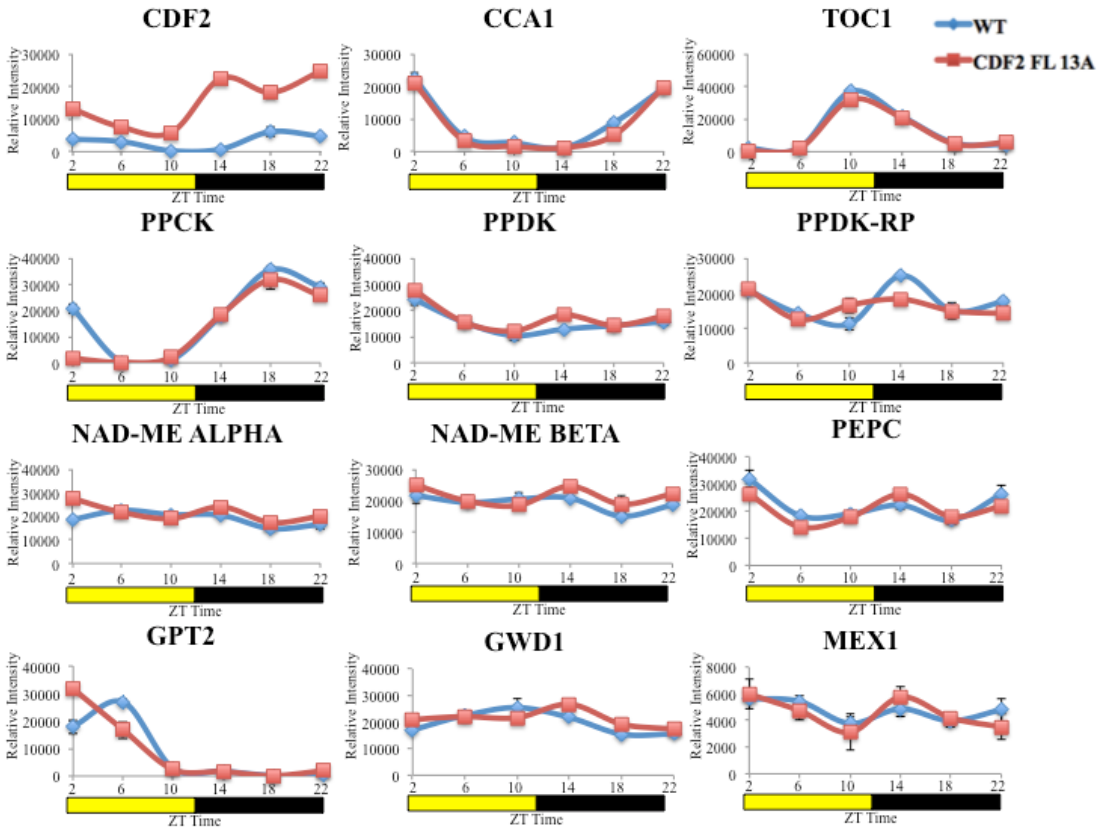

Figure 6.4. Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCDF2_FL_13A* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line FL_13A. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

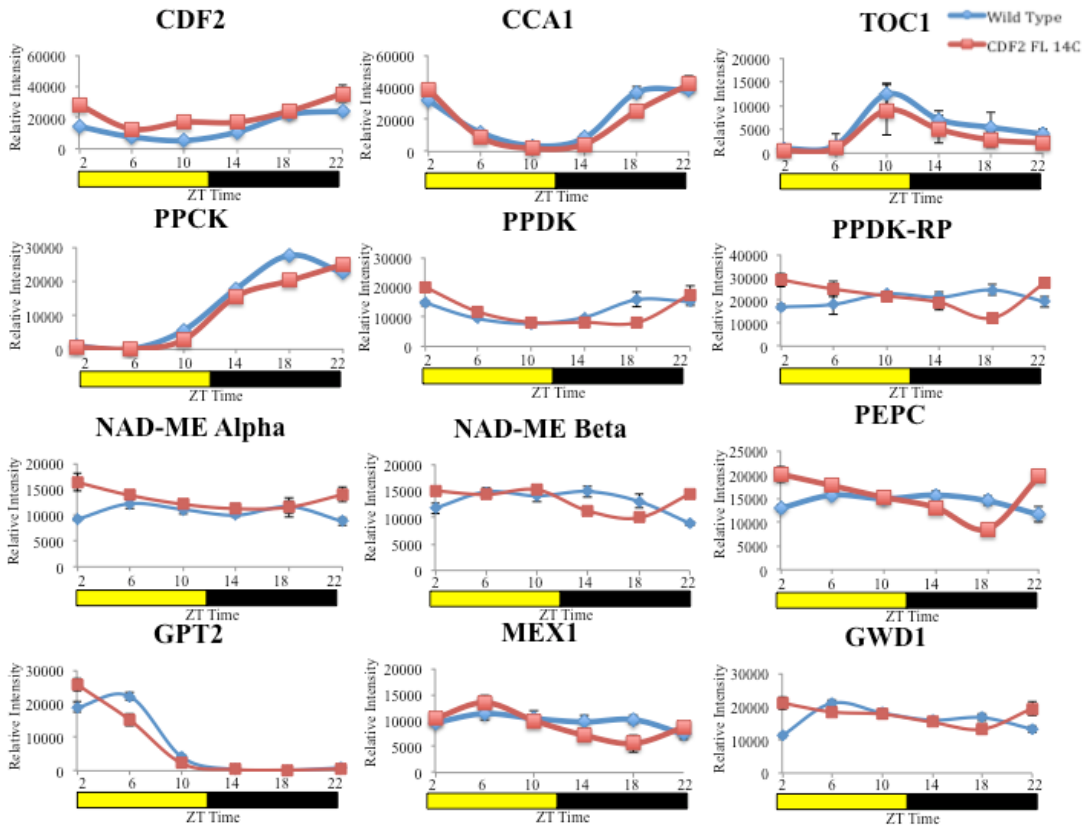

Figure 6.5. Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCDF2_FL_14C* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line FL_14C. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

6.2.2.2 Light/dark regulation of CAM and Clock genes in *KfCDF2* RNAi knockdown lines

For the RNAi lines *19A* and *26B*, clear differences were detected in the efficiency of the gene silencing effect of the RNAi construct that was designed to target the degradation of the endogenous *KfCDF2* transcripts in these lines. *KfCDF2_RNAi_19A* showed a 1.3-fold average decrease in the light and a 1.2-fold average decrease in the dark, whilst line *26B* showed a larger average decrease of 2.2- and 3.2-fold for the light and dark periods, respectively. In chapter 3.4.2, *KfCDF2_RNAi_26B* was also shown to be the more strongly silenced RNAi line. This suggested that these two lines had the potential to reveal dosage-dependent effects for *KfCDF2*. Furthermore, both RNAi lines showed similar trends in the regulation of the studied CAM and clock genes. Both showed changes in *NAD-ME, PEPC, PPDK-RP, MEX1* and *GWD*. For *NAD-ME α* and *β, PPDK-RP* and *GWD,* an increase in transcript abundance was only detected during the light period in line *RNAi_19A*, whereas in line *RNAi_26B* there was a consistent increase in the transcript abundance of these genes throughout the 24 h light/ dark cycle (Figs. 6.6 & 6.7). These results in themselves support the proposal of dosage-dependent effects of *KfCDF2* mis-expression, with line *26B* having the stronger RNAi effect. *KfPPCK1* and *KfPPDK* transcript levels were unaffected by the reduction in *KfCDF2* in either of the lines. Clock gene *KfCCA1* was also unaffected by *KfCDF2* down-regulation, whereas *KfTOC1* did display some small perturbations in both lines. In *RNAi_19A, KfTOC1* peaked at 10:00, with significantly higher transcript abundance than the wild type (Fig. 6.6), whereas this increase in expression was less noticeable in line *RNAi_26B*. In *KfCDF2_FL_19A*, *KfMEX1* showed a decrease in expression throughout the dark period, whilst in *KfCDF2_FL_26B KfMEX1* showed an increase. The daily peak of

KfGPT2 transcript abundance 2 h into the light period was phase delayed by 4 h in line *19A,* whereas in line *26B*, *KfGPT2* levels peaked at 02:00 in the light just as they did in the wild type.

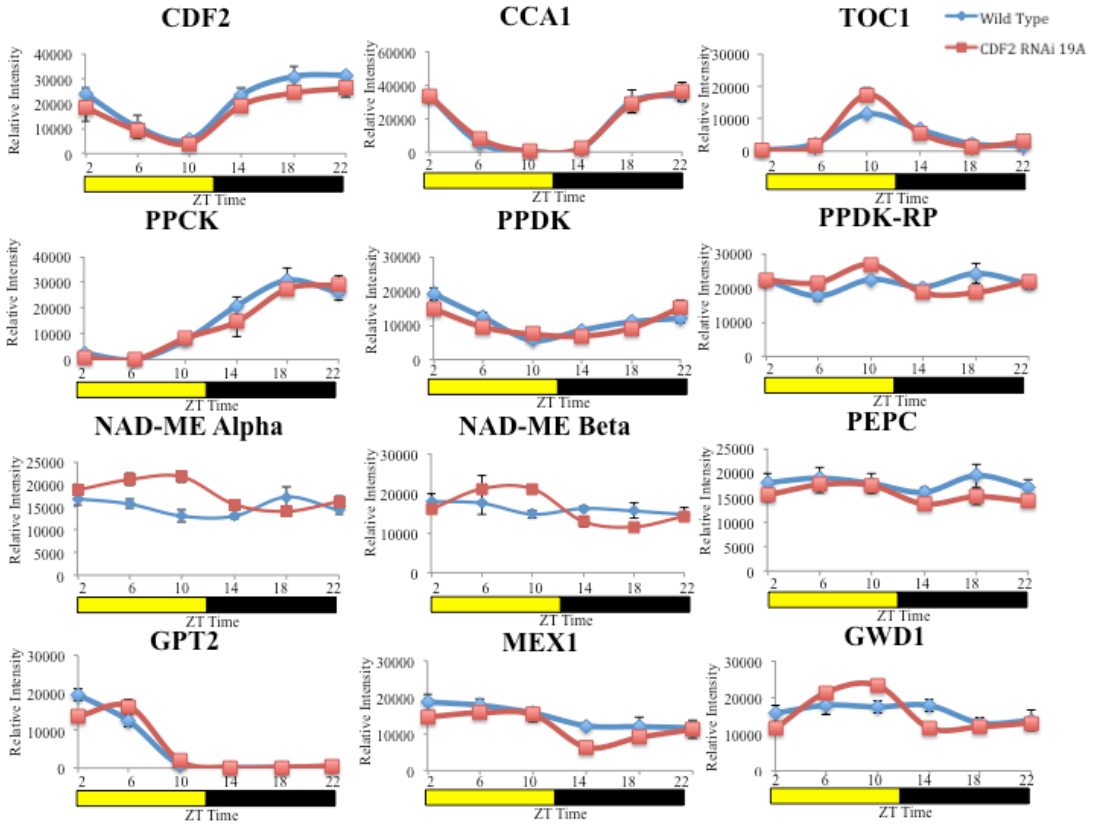

Figure 6.6. Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCDF2_RNAi_19A* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line RNAi_19A. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

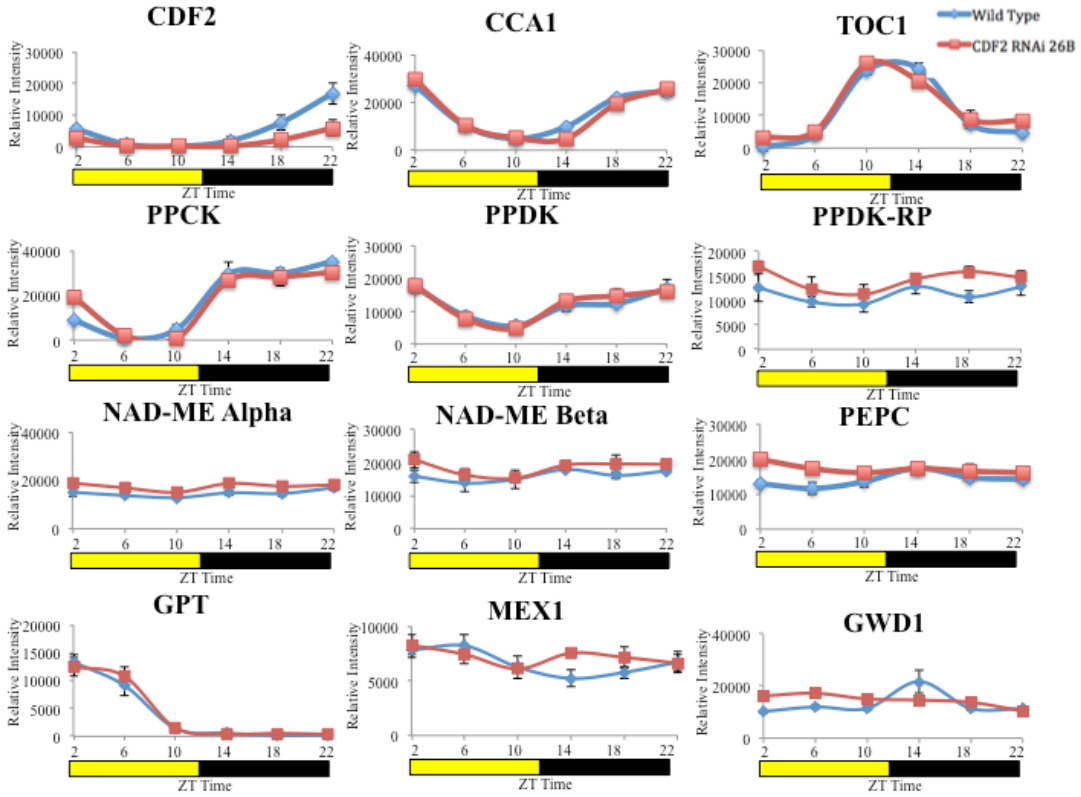

Figure 6.7. Light/ dark time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCDF2_RNAi_26B* (red) measured in full CAM leaves (leaf pair 6) sampled at 4 h intervals over a 12:12 light/dark cycle. Daily fluctuations in the transcript abundance of several clock genes (*KfCCA1 and KfTOC1*), and CAM genes (*KfPPCK, KfPPDK, KfPPDK-RP, KfNAD-ME α, KfNAD-ME β, KfPEPC, KfGPT2, KfMEX1* and *KfGWD1*) were confirmed for the wild type and mostly maintained in line RNAi_26B. The yellow bar below the graphs indicates the light period, and the black bar represents the dark period. The time course started at lights on at ZT0 (00:00 h), the first sample was collected at 02:00 and the growth chamber lights went off at 12:00.

6.2.3 The impact of perturbing *KfCDF2* levels on circadian oscillations in the transcript abundance of circadian clock controlled genes

It was shown in Chapter 3 that *KfCDF2* was under robust circadian clock control. It was therefore important to investigate whether this gene may play a role in the circadian optimisation of CAM. Transcript oscillations in the abundance of clock (*KfCCA1* and *KfTOC1*) and clock-controlled CAM-associated (*KfPPCK1*) genes were measured under LL conditions for leaf pair 6 sampled from both the wild type and the transgenic lines. For this experiment, one over-expressing line: *14C,* and one RNAi lines: *26B* was used.

6.2.3.1 Constant light free-running oscillations in the transcript abundance of CAM- and Clock-associated genes in a transgenic *K. fedtschenkoi KfCDF2* full-length over-expresser line

For *KfCDF2_FL_14C,* plants were entrained in 12:12 light/dark cycles and were then sampled on the 7th day at 10:00 (10 h into the 12 h light period) and 22:00 (10 h into the 12 h dark period). The light, temperature and humidity regime was then changed to constant light, 15˚C, 70 % relative humidity. Leaf pair 6 were sampled in biological triplicate starting at 02:00 LL, samples were then collected every 4 h for the next 3 days. Over-expression of *KfCDF2* was sustained throughout the time course in both LD and LL conditions (Fig. 6.8). After approximately 26 h under LL conditions, *KfCDF2* expression, which was still rhythmic at this time, peaked 4 h earlier than it did in the wild type, meaning *KfCDF2* peaked at the start of the light period (26:00 LL; Fig 6.8 – first black dotted box), whereas the wild type peaked in the middle of the second subjective light period (30:00 LL; Fig. 6.8). This phase advance in the *FL_14C* line held for the second subjective day in LL. After 60:00 LL (last subjective

light period in LL, third white dotted box) the *KfCDF2* rhythm in the *FL_14C* line dampened more rapidly than the dampening of the rhythm in the wild type, such that the transcript levels in both lines showed similar levels for the final few time points of the LL time course. It appeared that the over-expression of *KfCDF2* from the 35S promoter had in some way been nullified in the *14C* transgenic line towards the end of the LL time course, as *KfCDF2* was no longer being over-expressed relative to the WT (Fig. 6.8).

As with *KfCDF2, KfCCA1* and *KfTOC1* transcript levels peaked 4 h earlier than wild type for two of the three full oscillations in LL, and also showed slightly lower transcript levels at many of the LL time points (Fig. 6.8). *KfPPCK1* only showed a change in the timing of its transcript peak during the first 24 h in LL. After that it peaked in phase with the wild type (Fig. 6.8). However, as observed for a number of LL time points with *KfCCA1* and *KfTOC1* levels, *KfPPCK1* showed lower transcript abundance at many time points during the LL time course (Fig. 6.8). In particular, *KfPPCK1* levels were lower than the wild type at each of the transcript peaks during the LL conditions (Fig. 6.8).

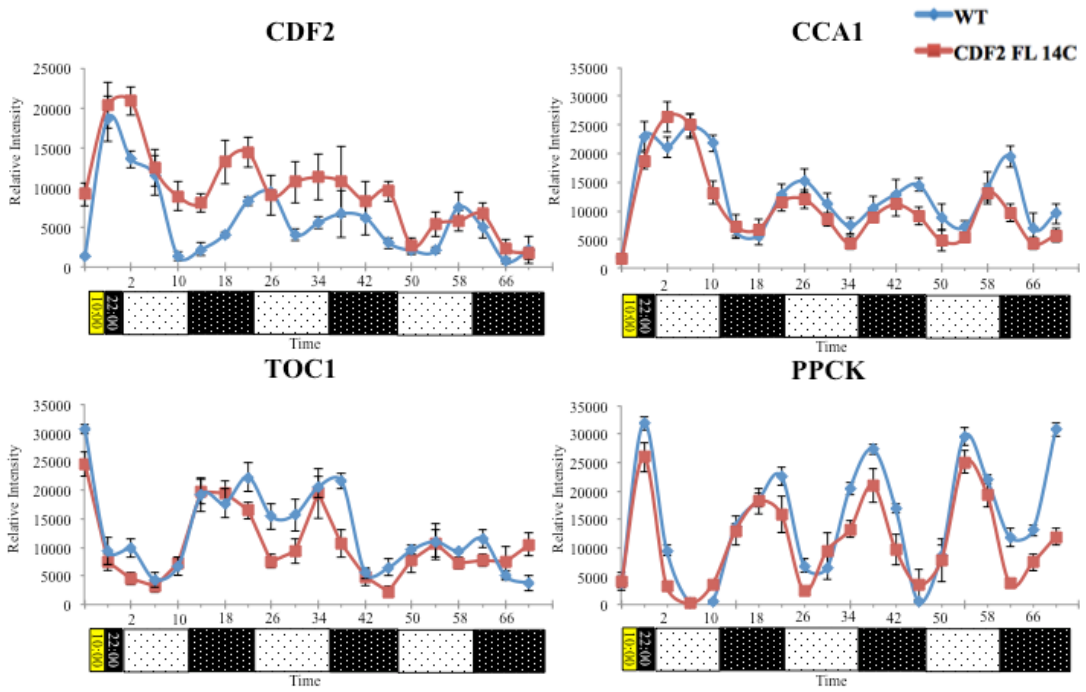

Figure 6.8. Constant light time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCDF2_FL_14C* (red) measured in full CAM leaves (LP6) sampled at 4 h intervals over a LL light regime. The first two samples collected were in LD at 10:00 L and 22:00 D, and then the light, temperature and humidity of the growth cabinet were set to constant levels (LL). Leaf pair 6 samples were collected every 4 h starting at 02:00 (2 h into the first subjective light period) in LL. Fluctuations in the transcript abundance of the clock genes *KfCCA1* and *KfTOC1*, and the CAM gene *KfPPCK1* were determined using semi-quantitative RT-PCR, along with *KfCDF2* expression, for both wild type and *KfCDF2_FL_14C*. All values were normalised to the abundance of the *KfUBQ10* reference gene. The yellow bar represents light period, the black bar represents the dark period. White boxes containing black dots represent subjective light and black boxes with white dots represent subjective dark. The plants were entrained for 7-days to lights on at ZT0 (00:00 h), and lights off at 12:00h. LL samples were collected from 02:00 in LL every 4 h for three days.

6.2.3.2 Constant light oscillations in CAM- and Clock genes from transgenic *KfCDF2* RNAi knockdown lines

The *KfCDF2_RNAi_26B* plants were entrained exactly as described above (4.2.3), but RNAi lines were left for 26h in LL before the first sample in constant conditions was collected, to enable lines to be in true free running conditions.

The RNAi knockdown phenotype of *KfCDF2* itself was maintained throughout the majority of the LL time course, and impacted on the plant's ability to maintain a rhythm in the abundance of *KfCDF2* transcripts (Fig. 6.9). Between 26:00 LL to 50:00 LL (Fig. 6.9; first white and black dotted boxes) there was little change in transcript abundance for *KfCDF2, KfCCA1* and *KfTOC1* in *KfCDF2_RNAi_26B*, whilst a robust oscillation could be seen in WT. After 50:00 LL, a rhythm in transcript abundance was detected for all three genes, but at a much lower amplitude than wild type, especially for *KfCCA1* and *KfTOC1*, where in *RNAi_26B* the expression is barely rhythmic. *KfCCA1* peaked 4 h later than the wild type (Fig. 6.9; second white and black dotted boxes), and *KfTOC1* also peaked after the wild type peak several times during the LL period, suggesting that the knockdown of *KfCDF2* leads to a phase delay in expression. The same thing also occurs with *KfPPCK* where the phase of gene expression appears to be delayed.

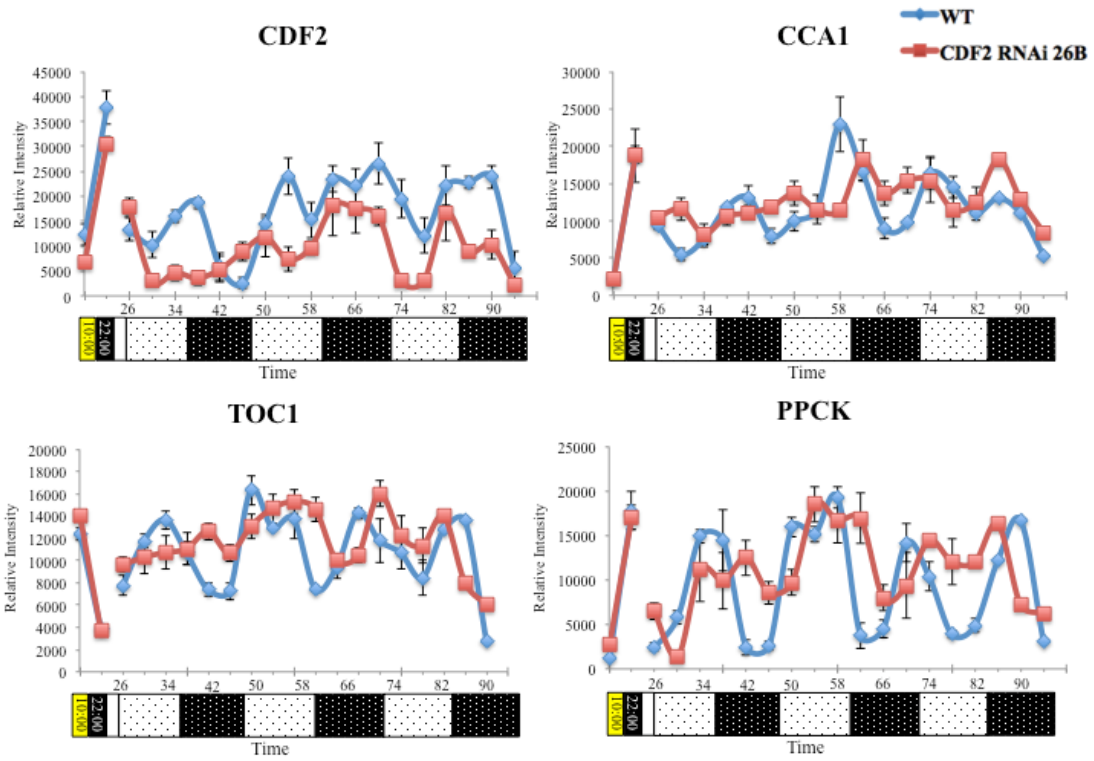

Figure 6.9. Constant light time course of the steady state transcript abundance of CAM- and circadian clock-associated genes in wild type *K. fedtschenkoi* (blue) and transgenic line *KfCDF2_RNAi_26C* (red) measured in full CAM leaves (LP6) sampled at 4 h intervals over a LL light regime. The first two samples collected were in LD and then the light regime was turned to LL. Plants were left for 26h in LL before sampling commenced. Fluctuations in the transcript abundance of clock genes (*KfCCA1* and *KfTOC1*) and CAM gene *KfPPCK* were determined using semi-quantitative RT-PCR, along with *KfCDF2* expression, for both Wild type and *KfCDF2_RNAi_26B*. The yellow bar represents light period, the black bar represents the dark period, and the white bar represents a break in sampling for the first 26h in LL. White boxes containing black dots represent subjective light and black boxes with white dots represent subjective dark. All values were normalised relative to the abundance of the KfUBQ10 reference gene. To entrain *KfCDF2_RNAi_26B*, The plants were entrained to lights on at ZT0 (00:00 h), and went off at 12:00h. Light sample was collected at 10:00 and dark sample was collected at 22:00. LL samples were collected from 26:00 in LL every 4 h for three days.

6.2.4 The impact of *KfCDF2* over-expression and RNAi in transgenic *K. fedtschenkoi* on the diurnal and circadian rhythms of CAM-associated CO_2 exchange

Both the over-expresser and RNAi *KfCDF2* transgenic lines had their gas exchange rhythms measured using the 6-channel, multi-cuvette IRGA system. Leaf pair 6 samples from developmentally synchronised plants were collected after 7-days entrainment under 12:12 light/dark conditions. The petioles of the leaves were submerged in water in small beakers that fitted into the gas exchange cuvettes. Gas exchange measurements were recorded for at least 5-days under 12:12 LD cycles before switching the light, temperature and humidity regime to LL conditions. The first 24 h of the gas exchange data was often very noisy as the leaves and the gas exchange system settle down, and so this part of the experimental data is not shown.

6.2.4.1 Impact of over-expressing *KfCDF2* in transgenic *K. fedtschenkoi* on diurnal gas exchange rhythms under LD cycles

Under both 12:12 LD and the subsequent LL conditions, both of the studied *KfCDF2_FL* over-expresser lines showed changes in their CO_2 exchange patterns relative to the wild type (Figs. 6.10 and 6.11). Under LD, both *KfCDF2_FL* lines showed greatly reduced phase IV in the light period, resulting in *KfCDF2_FL_13A* having a 99% reduction, and *KfCDF2_FL_14C* having an 83% reduction in total light period CO_2 fixation compared to the wild type (Table 6.3). For both lines, the wild type's large peak of nocturnal CO_2 fixation no longer occurred at the same time, with both lines showing a delay in CO_2 fixation relative to the wild type (Figs 6.10B and 6.11B). Line *13A* showed a phase shift in its daily CO_2 fixation peak, yet fixed large amounts of CO_2 for a longer time than wild type, resulting in an average increase in

dark CO_2 fixation over the four LD days of 35 % (Fig. 6.10; Table 6.3). Line *14C* on the other hand did not achieve such a large peak of CO_2 fixation, leading to a 25 % reduction during the dark relative to the wild type (Fig. 6.11; Table 6.2).

After entering LL conditions, both lines showed major changes in the circadian control of CO_2 exchange relative to wild type, as both *KfCDF2* lines struggled to oscillate their CO_2 fixation robustly (Figs. 6.10 & 6.11). After 3 oscillations for *KfCDF2_FL_13A* (60h LL) and 5 for *KfCDF2_FL_14C* (84h LL), CO_2 fixation became arrhythmic in both transgenic FL over-expressor lines. The wild type meanwhile maintained robust oscillations of CO2 fixation under the LL conditions with the classic short period rhythm characteristic of CAM in LL (Figs. 6.10C & 6.11C) *KfCDF2_FL_13A* became arrhythmic in LL earlier than *14C*, which correlated with line *13A* having higher levels of over-expression of *KfCDF2*.

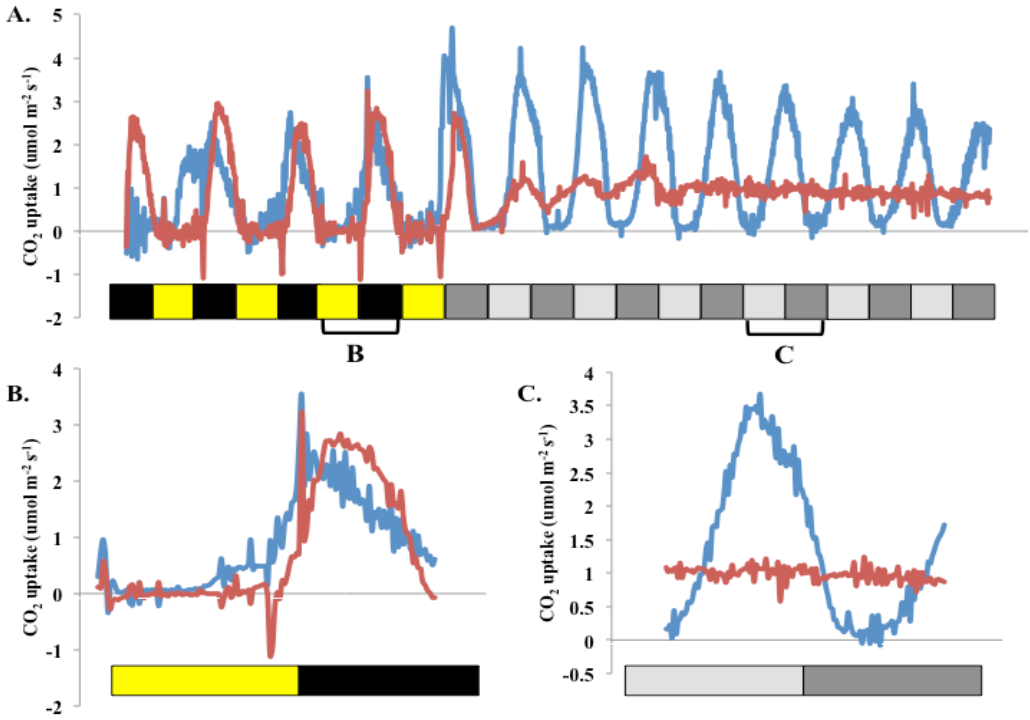

Figure 6.10. Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfCDF2_FL_13A*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfCDF2_FL_13A* is in red **A**. CO_2 fixation over 4 days LD followed by 7 days in LL. **B**. Close-up view of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C**. Close-up view of CO_2 fixation after entering LL, within 36 h the plant is clearly already displaying a shorter period than WT. The selected 24 h window is indicated by the C symbol below graph A.

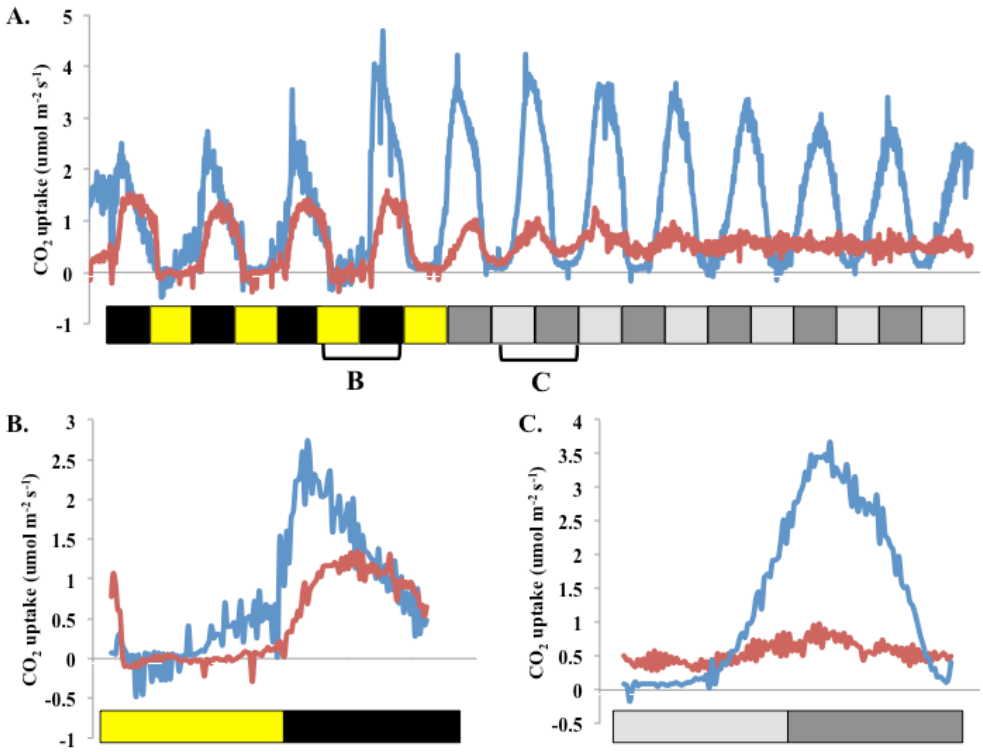

Figure 6.11. Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfCDF2_FL_14C*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfCDF2_FL_14C* is in red **A**. CO_2 fixation over 4 days LD followed by 7 days in LL. **B**. Snapshot of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C**. CO_2 fixation after entering LL, within 36h the plant is clearly already displaying a shorter period than WT. The selected 24 h window is indicated by the C symbol below graph A.

Table 6.2. Average net 24 h CO_2 fixation for each transgenic line and the WT over the entire light or dark period under LD conditions. Both transgenic lines fixed more CO_2 than the WT in the light, but less in the dark.

Line	Total CO_2 fixation over the 12 h light period (µmol CO_2 fixed $m^{-2)}$	% change	Total CO_2 fixation over the 12 h dark period (µmol CO_2 fixed $m^{-2)}$	% change
WT	18132.65	-	42652.85	-
CDF2 FL 13A	143.56	-99	57473.80	35
CDF2 FL 14C	3034.77	-83	32023.08	-25

6.2.4.2 Impact of silencing *KfCDF2* using RNAi in transgenic *K. fedtschenkoi* on diurnal gas exchange rhythms under LD cycles

During both 12:12 LD and LL conditions, *KfCDF2_RNAi_19A* showed a very similar pattern of gas exchange to the wild type; with the only difference being slightly more CO_2 was fixed (Fig. 6.12; Table 6.3). Closer inspection of the 24 h LD rhythm for line *19A* revealed that CO_2 fixation started earlier in the dark period in line *19A* compared to the wild type (Fig. 6.12B). This resulted in a 22 % increase in total dark CO_2 fixation (Table 6.4). Line *19A* also showed a more pronounced phase II, and also achieved a small amount of phase IV CO_2 fixation at the end of the light period. These changes also resulted in line *19A* achieving a slightly greater level of total light period CO_2 fixation by 6 % (Table 6.4). This trend was also maintained under LL conditions, where *19A* started fixing CO_2 slightly earlier than wild type for each oscillation, revealing a consistent phase advance in the oscillations of CO_2 fixation in this line (Fig. 6.12C), but collapsing into arrhythmia did not occur. This may be due to 19A only having small decreases in expression level, which may not severely affect the clock.

Line *KfCDF2_RNAi_26B* did not show a phase IV gas exchange signature, and also did not begin to fix CO_2 until later in the dark period than either the wild type or line *19A* (Fig. 6.13). Furthermore, line *KfCDF2_RNAi_26B* also fixed much less CO_2 than the wild type during the dark period, with an overall reduction in total dark period CO_2 fixation of 26 % (Table 6.4; Fig 6.13B). During the light period, the total amount of CO_2 fixation was similar for both the wild type and line *26B* (Table 6.4; Fig. 6.13B).

Under LL free-running constant conditions, *KfCDF2_RNAi_26B* showed a dampening of the circadian rhythm of CO_2 exchange such that after 120 h LL (Fig.

6.13; the last two grey shaded boxes) the gas exchange pattern had become arrhythmic, whilst the wild type rhythm of CO_2 fixation was still oscillating robustly (Fig. 6.13C).

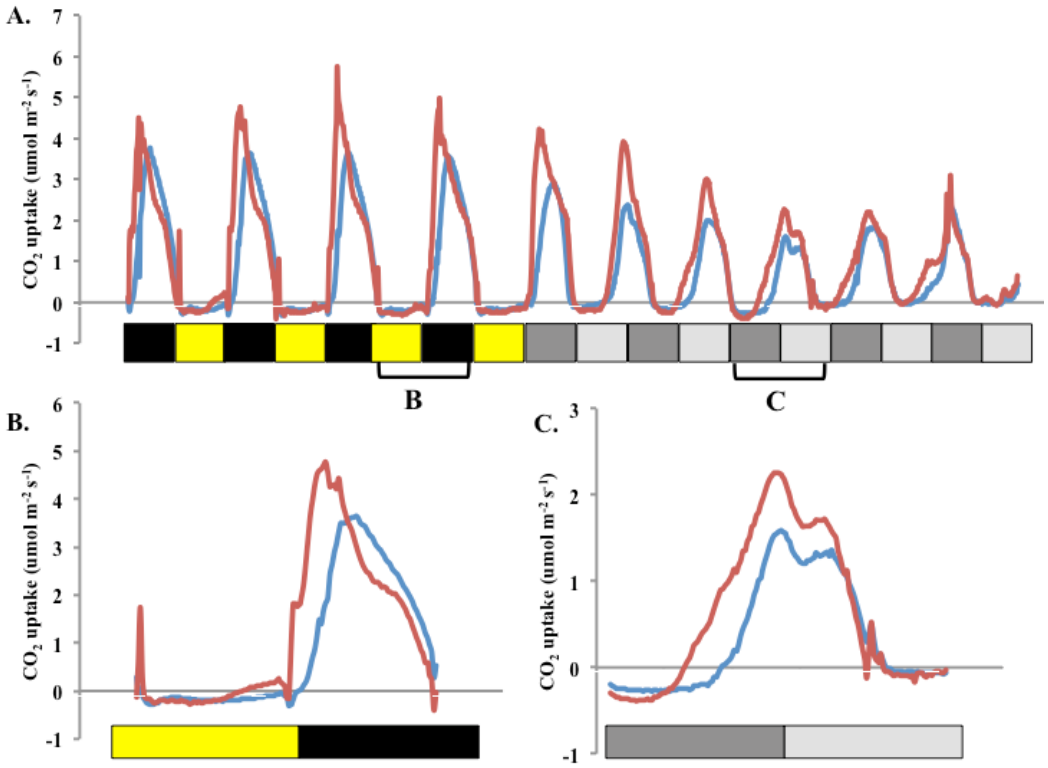

Figure 6.12. Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfCDF2_RNAi_19A*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfCDF2_RNAi_19A* is in red **A**. CO_2 fixation over 4 days LD followed by 7 days in LL. **B**. Close-up view of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C**. Close-up view of CO_2 fixation after entering LL,. The selected 24 h window is indicated by the C symbol below graph A.

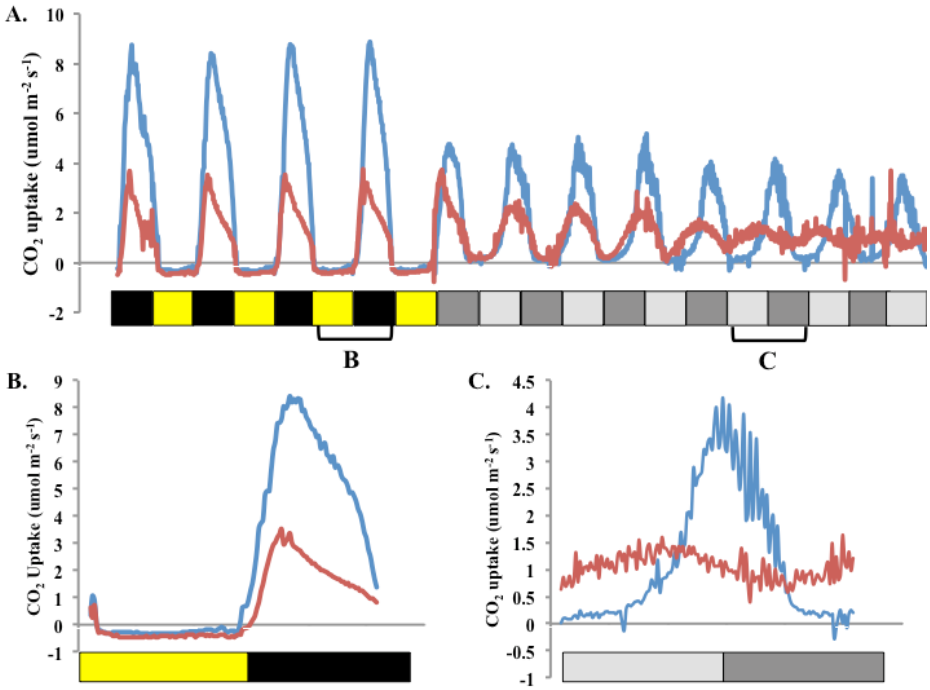

Figure 6.13. Cycles and rhythms of CO_2 exchange in LD and LL conditions for *KfCDF2_RNAi_26B*. Black boxes represent lights off, yellow boxes show lights on. The dark grey boxes show subjective dark and the light grey boxes show subjective light once the plant enters LL conditions. WT is in blue and *KfCDF2_RNAi_26B* is in red **A**. CO_2 fixation over 4 days LD followed by 7 days in LL. **B**. Snapshot of CO_2 fixation over a 24 h LD period. The selected 24 h period is indicated by the B symbol below graph A. **C**. CO_2 fixation after entering LL. The selected 24 h window is indicated by the C symbol below graph A.

Table 6.3. Average net 12 h CO_2 fixation for each transgenic line and the WT over the entire light or dark period. Both transgenic lines fixed more CO_2 than the WT in the light, but less in the dark.

	Total CO_2 fixation over the 12 h light period (µmol CO_2 fixed $m^{-2)}$	% change	Total CO_2 fixation over the 12 h dark period (µmol CO_2 fixed $m^{-2)}$	% change
WT	-6979.67	-	86089.57	-
CDF2 RNAi 19A	-6557.42	6	105413.07	22
WT	2810.27	-	61801.68	-
CDF2 RNAi 26B	2736.07	-0.02	45425.80	-26

6.2.5 Measurements of the end products of CAM using metabolite assays

6.2.5.1 Impact of overexpressing *KfCDF2* on major pools of primary metabolites associated with CAM

6.2.5.1.1 Daily oscillations in leaf malate content

In chapter 3, malate levels at dawn and dusk were estimated based on the colour change of the pH sensitive dye chlorophenol red. This assay was rather crude and not robustly quantitative, but was a powerful method for rapidly screening through the hundreds of transgenic lines that were available for screening at the start of this project. Having identified the potentially interesting transgenic lines using the experiments described in Chapter 3, it was important to make quantitative measurements of the level of malate in the leaves of each line using enzyme-linked spectrophotometric assays for malate (using the method described by Möllering, 1974).

The wild type and both over-expresser lines displayed significant fluctuations in malate between dawn, when there was high malate (Fig. 6.14 – (D)ark bars), and dusk when there was low malate (Fig. 6.14 – (L)ight bars). For the wild type, the Δ-malate value was 40.1 μmol gFW^{-1}, whilst the value for *KfCDF2_FL_13A* was 17.2 μmol gFW^{-1}, and for *KfCDF2_FL_14C* the malate differential was 16.3 μmol gFW^{-1}. Thus, both *KfCDF2* over-expresser lines showed decreased malate accumulation and daily turnover. These changes in malate accumulation during the dark period were found to be statistically significant (Dark 13A: $P < 0.0001$, Dark 14C: $P = 0.0002$). Both over-expresser lines also showed increased malate retention at the end of the light relative to wild type, but due to larger variation between the biological replicates in the light, this could not be confirmed as statistically significant (Fig. 6.14).

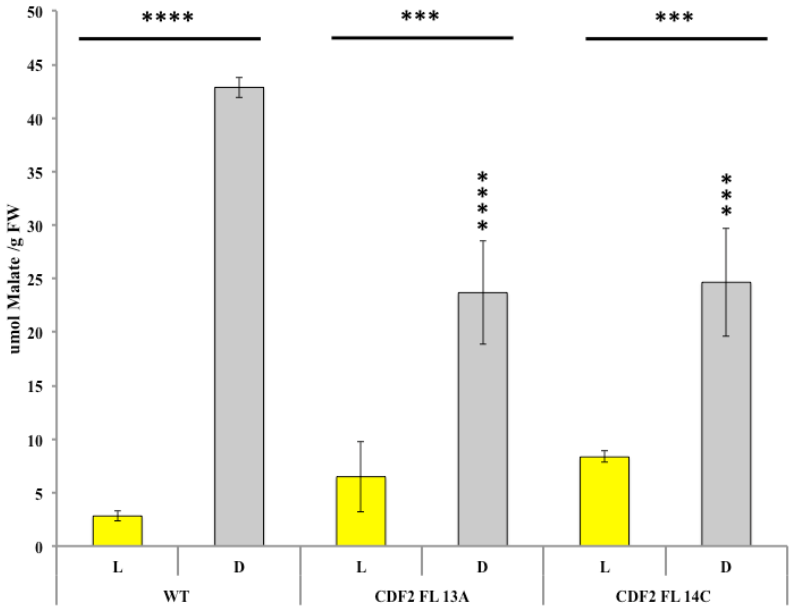

Figure 6.14. Diurnal fluctuations in malate concentrations in *KfCDF2_FL_13A* and *KfCDF2_FL_14C*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Malate levels were low at the end of the light period due to malate decarboxylation occurring during the day, and high at the end of the dark due to nocturnal CO_2 fixation resulting in malic acid accumulation. Malate levels for *KfCDF2_FL* lines were slightly higher than WT at the end of the light period and malate levels were lower at the end of the dark. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one bar represent significant differences relative to the corresponding WT value.

6.2.5.1.2 Daily fluctuations in leaf starch

As malate and starch levels reciprocate over each 24 h period during CAM, it was also of importance to quantify starch levels in both wild type and transgenic lines. In Chapter 3, starch levels were only approximated using iodine staining of cleared leaf discs.

Relative to the wild type, both over-expression lines showed lower levels of starch at the end of the dark period (WT = 1.4 mg Starch gFW^{-1}, 13A = 0.9 mg Starch gFW^{-1}, 14C = 0.8mg Starch gFW^{-1}), although neither were shown to have significantly lower

amounts due to wild type starch levels showing large levels of variation between the biological replicates. Over-expression of *KfCDF2* appears led to an increase in starch turnover. The wild type had an average Δ-starch (the differential between the level of starch at dusk and dawn) of 1 mg gFW^{-1}. *FL_13A* had a Δ-starch of 1.5 mg gFW^{-1} and *FL_14C* turned over 2.4 mg gFW^{-1} within the 24 h cycle. Due to large variations between the individual biological replicates, these results were not found to be statistically significant from one another.

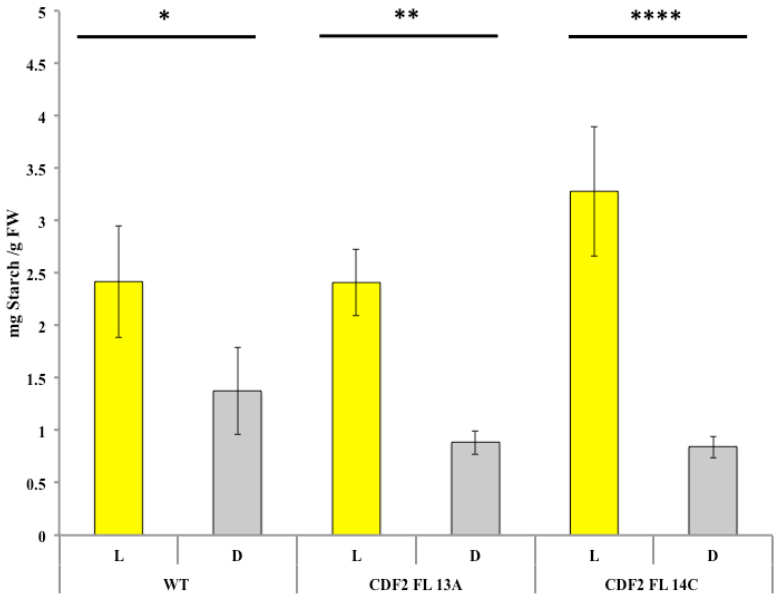

Figure 6.15. Diurnal fluctuations in starch concentrations in wild type, *KfCDF2_FL_13A* and *KfCDF2_FL_14C.* Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Starch levels are highest at the end of the light period, after photosynthesis has been occurring all day. Starch levels then drop to lower levels at the end of the dark, due to starch stores being mobilised to fuel nocturnal respiration and produce PEP for nocturnal CO_2 fixation. Both FL lines showed a reduction in mean starch levels during the dark period, whilst only *KfCDF2_FL_14C* also showed an increased level of starch in the light. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one bar represent significant differences to the corresponding WT value.

6.2.5.1.3 Daily fluctuations in soluble sugars

The results had revealed that the FL lines fixed less CO_2 than the wild type in the dark (Table 6.2), and accumulated less malate by dawn (Fig. 6.14). The same lines also displayed an increase in starch turnover (Fig. 6.15). These results suggest that the increase in starch turnover was not involved in the production of more PEP for dark CO_2 fixation, and so it was important to investigate the levels of soluble sugars in the leaves of these plants as a second possible product from the increased starch turnover. The soluble sugars assayed were glucose, fructose and sucrose (Fig. 6.16).

Both *KfCDF2_FL* lines showed a significant increase in the amount of glucose present, with glucose levels at the end of the light in line *14C* being approximately 15 times greater than glucose in the wild type (Fig. 6.16; 13A: $P = 0.0208$, 14C: $P < 0.0001$). By contrast, both lines showed similar fructose levels at the end of the light when compared to the wild type. *KfCDF2_FL_13A* also showed a statistically significant increase in its leaf sucrose level at the end of the light ($P = 0.0003$). Line *14C* also showed an increase in sucrose at the end of the light, but this was not significant.

At the end of the dark period, the two *FL* transgenic lines yielded contrasting results. *KfCDF2_FL_13A* showed no significant differences to wild type, although the average amount of both glucose and sucrose was higher than in wild type. The level of fructose also remained approximately the same as the wild type in line 13A at the end of the dark. However, line *KfCDF2_FL_14C* had a significant and very large increase in both fructose and sucrose levels compared to wild type at the end of the dark (*14C* fructose: $P < 0.0001$, *14C* sucrose: $P < 0.0001$). *14C* also showed a reduced level of glucose at the end of the dark period relative to the wild type, but the difference was not found to be significant.

KfCDF2_FL_14C turned over large quantities of soluble sugars each day, whereas *KfCDF2_FL_13A* did not show such dramatic significant differences between the dusk and dawn levels of each soluble sugar. Fructose levels in line 13A did not show any significant fluctuation between dusk and dawn (Fig. 6.16).

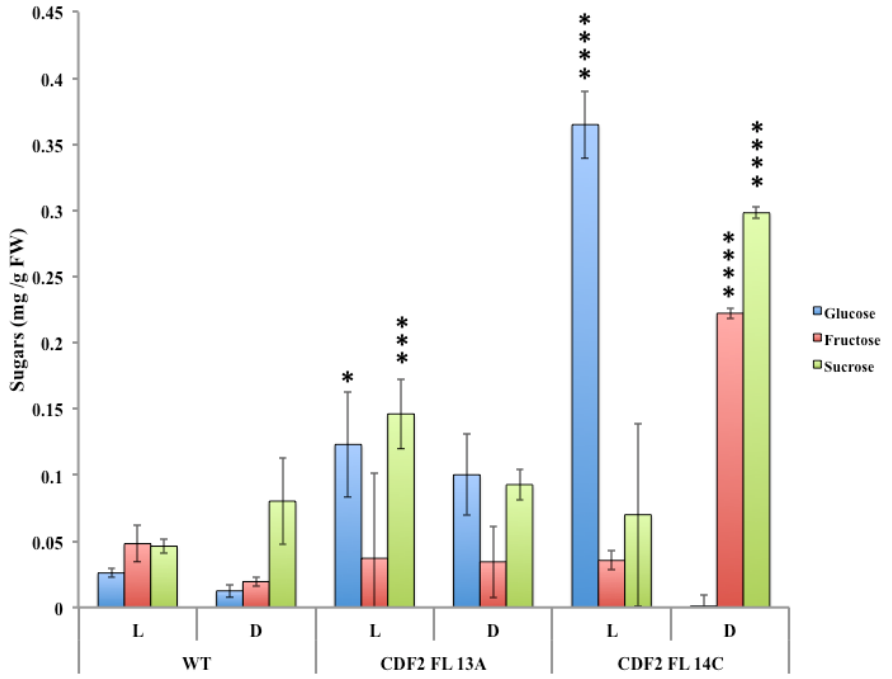

Figure 6.16. Levels of Glucose, Fructose and Sucrose in *KfCDF2* overexpressing lines at the end of the light (L; 23:00) and dark (D; 11:00) periods. *KfCDF2_FL_13A* shows increases in light levels of glucose and sucrose, and *KfCDF2_FL_14C* also shows high glucose at the end of the light. *KfCDF2_FL_14C* also shows an increase in both fructose and sucrose at the end of the dark period. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

Summing the total of all three sugars allowed a comparison to be made between the total amount of soluble sugars in the leaves of each line at the end of the light and the dark (Fig. 6.17). The two transgenic *FL* lines both showed an increase in the total amount of soluble sugars present in LP6 over the two time points compared to wild type (*13A* and *14C* compared to wild type: $P < 0.0001$), and *13A* and *14C* had similar total values. Both transgenic lines also had significant increases in the total quantity of sugar in the leaves in the light, with *14C* showing the largest increase (Fig. 6.17). *KfCDF2_FL_13A* also had an increase in total dark sugar levels compared to wild type ($P < 0.0001$), whereas *14C* showed no significant difference to wild type (Fig. 6.17). This therefore still does not explain where the extra dark period starch that was being turned over in both transgenic *FL* lines was going to. Once possible explanation would be that it was being used to fuelled extra PEP production, but there was perhaps an issue with the enzymes involved in CO_2 fixation or because changes in circadian clock control in *KfCDF2_FL* lines may lead to a failure to achieve correct circadian control of CAM, thus leading to the futile cycling, leading to malate being decarboxylated simultaneously as CO_2 is being fixed. This would require extra ATP and reducing equivalents, which in turn could only be fuelled by extra starch and soluble sugar turnover.

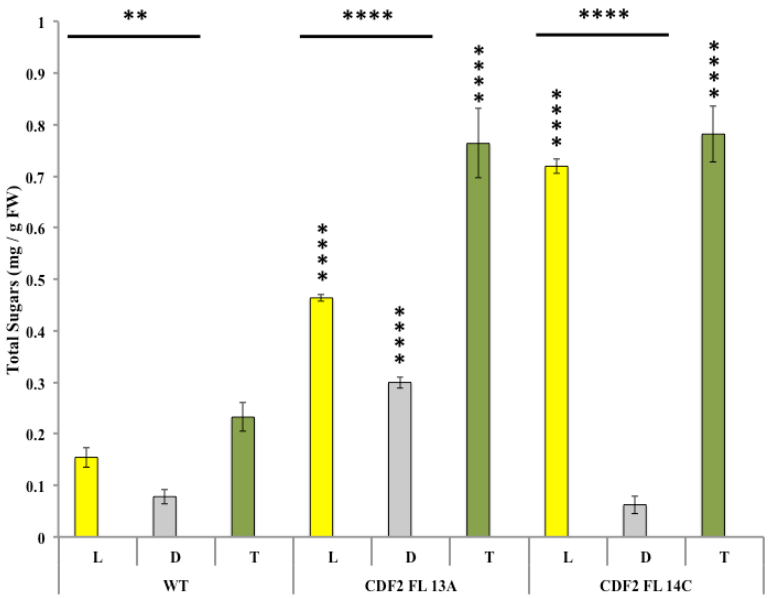

Figure 6.17. Variation in total soluble sugar levels in WT and the *KfCDF2 FL* lines. Yellow bars indicate light levels of sugars (L), grey bars represent dark levels (D), and green bars show the total levels of soluble sugars (T) calculated by summing the end-of-dark and end-of-light values for all sugars. There was a clear trend in that these *KfCDF2_FL* lines had higher sugar levels than the wild type during the light period. *KfCDF2_FL_13A* also had higher levels of total sugar during the dark, whilst *14C* showed no difference from the wild type in the dark. Both lines performed approximately equal levels of total sugar turnover over the 24 h period. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one bar represent significant differences to the corresponding WT value.

6.2.5.2 Impact of reducing the expression of *KfCDF2* in transgenic *K. fedtschenkoi* on the products of CAM

6.2.5.2.1 Daily fluctuations in malate

Both *KfCDF2_RNAi* lines and the wild type showed significant cycling of malate between dawn and dusk **(**Fig. 6.18, $P < 0.0001$**)**. The two RNAi lines contrasted differently relative to the wild type. *KfCDF2_RNAi_19A* showed a significant increase in malate levels in the light period ($P < 0.0001$) and during the dark ($P = 0.0067$). On the other hand, *KfCDF2_RNAi_26B* showed no significant differences to the wild type (Fig. 6.18). The Δ-malate values for each of the lines were 40.1 μmol gFW^{-1} for wild type, 18.7 μmol $gFW^{-1}$for *19A* and 41.8 μmol $gFW^{-1}$for *26B*. Thus, the wild type and line *RNAi_26B* turned over similar amounts of malate between dawn and dusk, whereas *KfCDF2_RNAi_19A* turned over significantly less (Fig. 6.18).

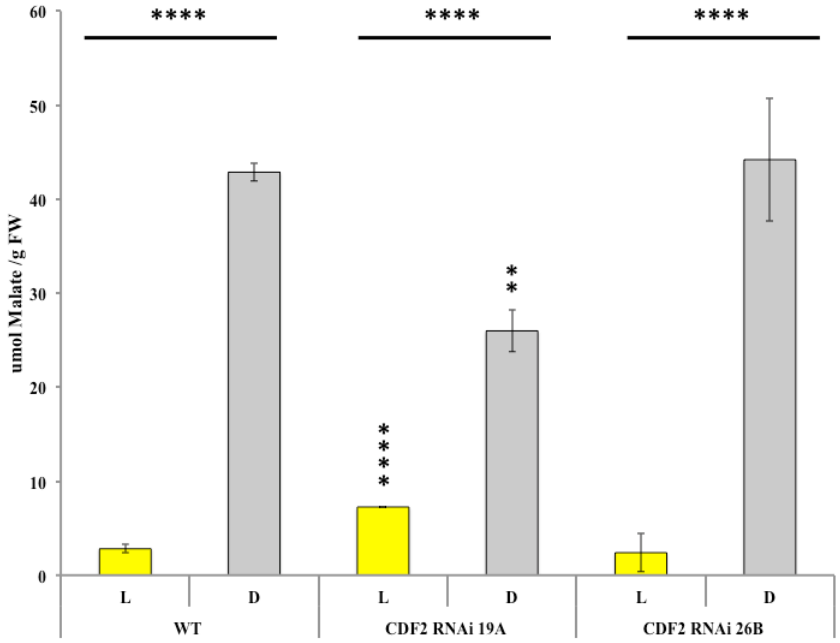

Figure 6.18. Diurnal fluctuations in malate concentrations in *KfCDF2_RNAi_19A* and *KfCDF2_RNAi_26B*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Malate levels were low at the end of the light period due to malate decarboxylation occurring during the day, and high at the end of the dark due to nocturnal CO_2 fixation resulting in malic acid accumulation. Malate levels for *KfCDF2_RNAi* lines were slightly higher than WT at the end of the light period, with *KfCDF2_RNAi_19A* being significantly so, and malate levels were also significantly lower for *KfCDF2_RNAi_19A* at the end of the dark. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $\leq$ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences relative to the corresponding WT value.

6.2.5.2.2 Daily fluctuations in starch

KfCDF2_RNAi_19A did not carry out significant cycling of starch over the 24 h period, whereas wild type and *KfCDF2_RNAi_26B* had more starch at the end of the light and less at the end of the dark (Fig. 6.19). *KfCDF2_RNAi_26B* turned over much more starch than the wild type (*26B* = 2.6 mg Starch gFW^{-1} turnover, wild type = 1 mg Starch gFW^{-1} turnover) due to there being a significantly higher amount of starch at the end of the light in line *26B* ($P = 0.0006$). *KfCDF2_RNAi_19A* turned over 0.5 mg Starch gFW^{-1}, but had slightly more starch present in the dark period relative to

the light period. The other two lines had more starch at the end of the light than at the end of the dark, consistent with the expected diurnal cycle of starch being synthesised as the product of photosynthesis and pyruvate recycling in the light, and degraded to fuel nocturnal respiration and PEP provision for PEPC-mediated CO_2 fixation in the dark (Fig. 6.19).

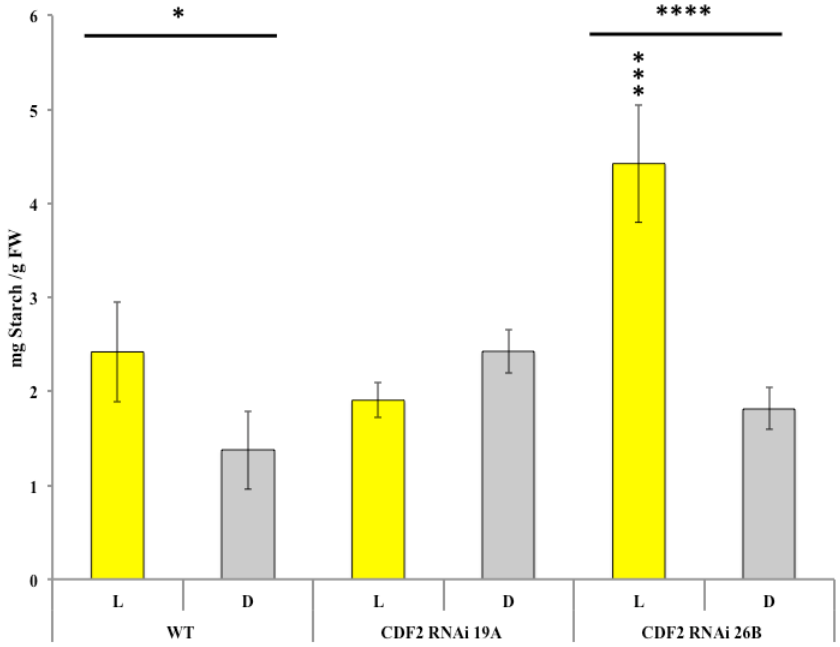

Figure 6.19. Diurnal fluctuations in starch concentrations in *KfCDF2 RNAi_19A* and *KfCDF2_RNAi_26B*. Values represent the mean of three biological replicates sampled at dawn (D, 23:00) and dusk (L, 11:00) in 12:12 light/dark conditions. Starch levels were highest at the end of the light period, as a result of accumulation due to photosynthesis occurring throughout the light period. Starch levels then dropped to lower levels at the end of the dark due to starch stores being mobilised to produce PEP for nocturnal CO_2 fixation and for respiration. Both RNAi lines showed a slight increase in dark starch stores, whilst *RNAi_26B* also shows a significant increase at the end of light. *RNAi_19A* possessed a slightly decreased level of starch at the end of the light. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = $\leq$ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one bar represent significant differences to the corresponding WT value.

6.2.5.2.3 Daily fluctuations in soluble sugars

Both transgenic lines only showed significant differences compared to the wild type at the end of the dark period (Fig. 6.20). The RNAi lines showed significant increases in glucose (*19A*: P = 0.0194, *26B*: P <0.0001), and sucrose (*19A*: P = 0.043, *26B*: P <0.0001). *KfCDF2_RNAi_19A* also showed a significant increase in fructose (P <0.0001), whereas although *26B* showed an increase, it was not significant due to high variance between the biological replicates. In the light, despite levels not being significantly different to the wild type, both lines did show a decrease in fructose compared to wild type (Fig. 6.20).

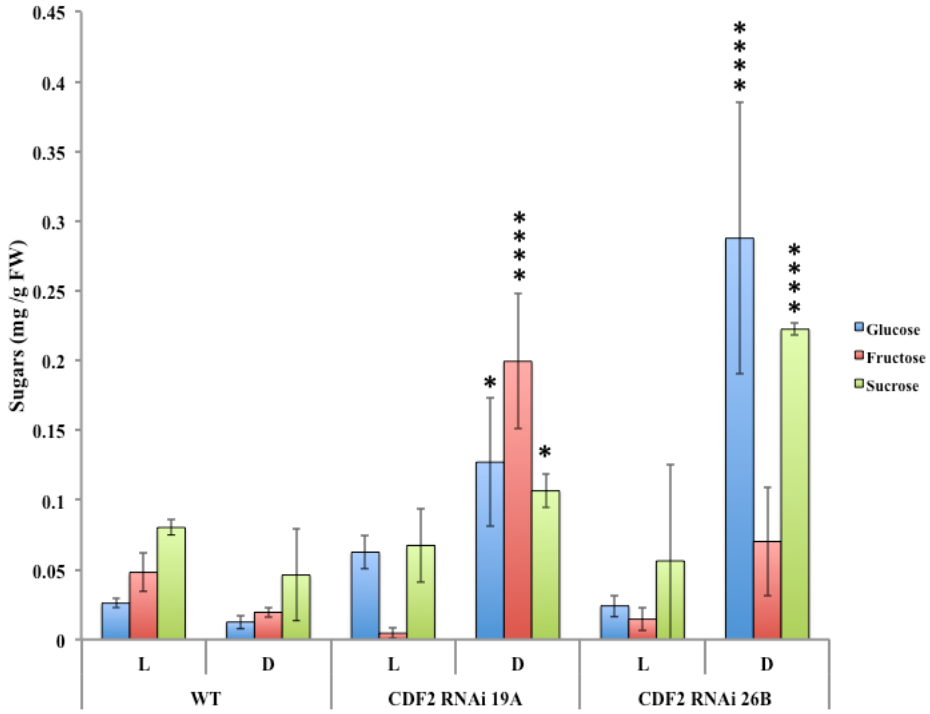

Figure 6.20. Levels of Glucose, Fructose and Sucrose in *KfCDF2_RNAi* lines, at the end of the light (L; 23:00) and dark (D; 11:00) periods. *KfCDF2_RNAi_19A* shows a significant increase in all three sugars assayed at the end of the dark, whilst *KfCDF2_RNAi_26B* also showed increased levels of all three sugars, only glucose and sucrose levels were significantly increased. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one bar represent significant differences to the corresponding WT value.

As with the overexpressing lines, both RNAi lines showed a general increase in total sugars throughout the whole 24 h period (Fig. 6.21 – T; *19A*: P = 0.008, *26B*: P = 0.006). Both lines were also found to possess increased levels of the total amount of soluble sugars found in leaves at the end of the dark period (*19A*: P = 0.048, *26B*: P = 0.041). Neither showed a significant difference to wild type at the end of the light period, but on average, they did have small decreases in total sugar quantity (Fig. 6.21).

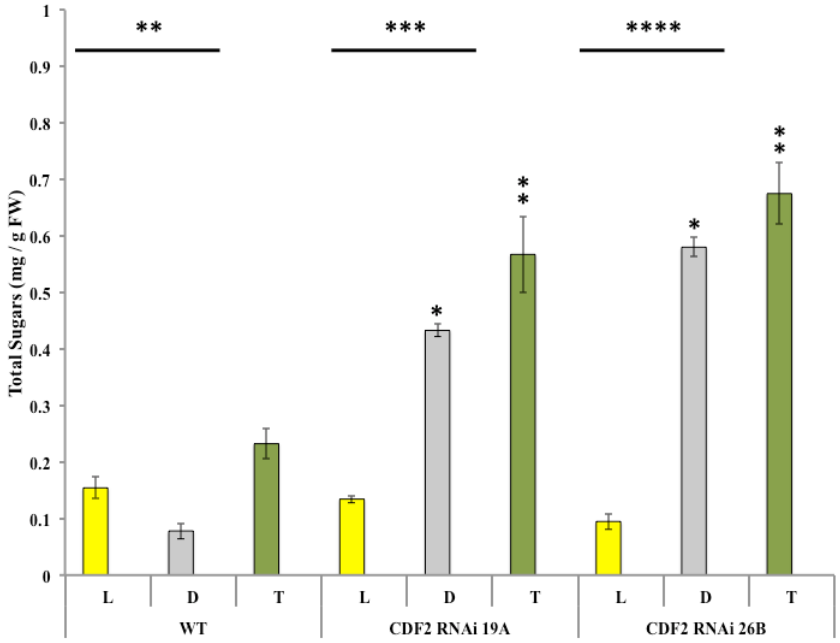

Figure 6.21. Variation in total soluble sugar levels in WT and the *KfCDF2_RNAi* lines. Yellow bars indicate light sugars levels (L), grey represent dark levels (D) and green bars show the total of sugar levels (T) calculated by summing the end-of-dark and end-of-light values for all sugars. There is a clear trend that these *KfCDF2_RNAi* lines have higher sugar levels during the dark period, leading to a total sugar increase over the 24h. Both lines show slight reductions in total sugar levels during the light period. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one bar represent significant differences to the corresponding WT value.

6.2.6. Immunoblot determination of the relative protein abundance of key enzymes associated with CAM

The *KfCDF2* transgenic lines were found to display a number of major changes in terms of their daily CO_2 exchange, especially nocturnal CO_2 fixation, Δ-malate, and starch and sugar turnover. However, the changes in the transcript abundance of the target gene were relatively small and the knock-on perturbations to other genes associated with CAM and the circadian clock were at best subtle. It was therefore important to see if the changes in CO_2 fixation and the associated major metabolic pools were a consequence of changes in PPCK protein abundance, or the phosphorylation state of its target protein PEPC, the primary nocturnal carboxylase during CAM. Samples were collected every 4 h over a 12:12 light/ dark time course, using the same experimental conditions used for the previous gene transcript level studies.

6.2.6.1 Impact of over-expressing *KfCDF2* on CAM protein abundance and function

Both *KfCDF2_FL* lines showed similar and constant levels of PEPC protein over the 12:12 light/ dark time course (Fig. 6.22A). Both lines also showed a decrease in PEPC phosphorylation state at the beginning of the dark period, suggesting that there was either less PPCK enzyme to phosphorylate PEPC, or PPCK was no longer working optimally (Fig. 6.22B). It could also suggest that PP2A, which de-phosphorylates PEPC, had increased in activity. This decrease in PEPC phosphorylation would be predicted to impact directly on gas exchange and malate accumulation, as had already been demonstrated in sections 6.2.4 and 6.2.5.

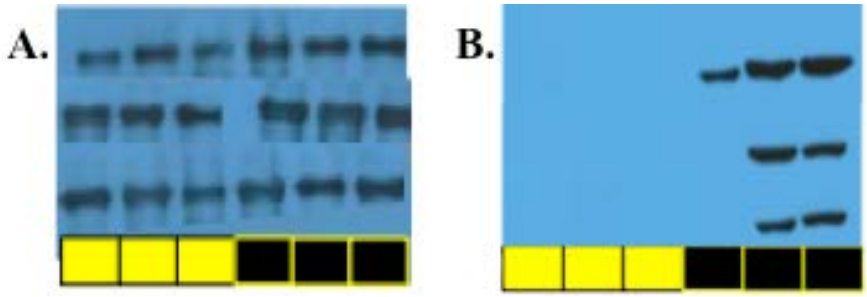

Figure 6.22. Immunoblotting revealed that PEPC protein levels do not change in abundance, but phosphorylation state does in the *KfCDF2_FL* overexpresser lines. **A.** Immunoblot for PEPC **B.** Immunoblot for phosphorylated PEPC; (Top –WT, Middle – *FL_13A*, Bottom – *FL_14C*).

6.2.6.1 Impact of reducing *KfCDF2* expression on protein abundance and function

Due to time constraints and problems extracting protein from *KfCDF2_RNAi_19A*, only protein levels for *KfCDF2_RNAi_26B* could be determined (Fig. 6.23). PEPC levels were again constant throughout the 12:12 light/ dark cycle, and were very similar between wild type and *26B* (Fig. 6.23A). The phopho-PEPC immunoblot revealed that the wild type had again begun to phosphorylate PEPC at 14:00 dark (2 h into the 12 h dark period), whereas the *RNAi_26B* line did not display a detectable level of PEPC phosphorylation at 14:00, and also appears to have slightly lower levels of phospho-PEPC at 18:00 and 22:00 (Fig. 6.23B). This may be the reason underlying the fact that this line and the two *FL* lines all showed delays in their rise in dark CO_2 fixation at the beginning of the dark. It would be interesting to see if *KfCDF2_RNAi_19A* showed a contrasting result for the phospho-PEPC blot, as that line showed an increase in its dark CO_2 fixation ahead of the wild type.

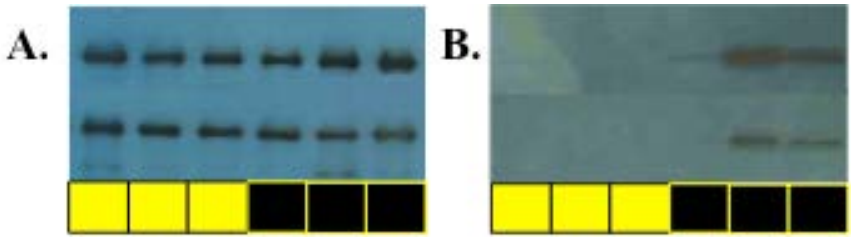

Figure 6.23 Immunoblotting revealed that PEPC protein levels do not change in abundance, but phosphorylation state does in the *KfCDF2_RNAi* lines. **A.** Immunoblot for PEPC **B.** Immunoblot for phosphorylated PEPC; (Top –WT, Bottom – 26B).

6.2.7 Total extractable PEPC activity and the apparent K_i of PEPC for malate in the *KfCDF2* lines

Whilst the immunoblots provided information about the abundance of PEPC and its phosphorylation state, it was important to investigate whether changes to *KfCDF2* expression led to alterations in the activity of PEPC, and/ or its malate sensitivity, which is the biochemical consequence of PPCK activity and PEPC phosphorylation. To quantify the total extractable activity of PEPC, it was expressed relative to both protein and chlorophyll in order to control for changes to *KfCDF2* expression affecting the abundance of either of these reference levels.

6.2.7.1 Enzyme assay results for *KfCDF2* full-length over-expression lines

6.2.7.1.1 Total extractable activity of PEPC

Both *KfCDF2_FL* lines showed a significant increase in total extractable activity of PEPC per mg of protein (Fig. 6.24; *13A*: P <0.0001, *14C*: P = 0.0018). *KfCDF2_FL_13A* showed a huge 44.8-fold increase, and *14C* showed a 14.1-fold increase in activity.

When PEPC activity was expressed per mg of chlorophyll though, neither line showed a significant increase, despite them both showing a slight average increase in

activity (Fig. 6.24**,** *13A*: 1.5 fold increase, *14C*: 1.6 fold increase in activity). The chlorophyll results fit with the immunoblot data more, as we did not see a change in PEPC protein abundance, and so it is difficult to understand why line *13A* had a 44.8-fold increase in total PEPC activity when expressed per mg of protein. However, the explanation could simply be that the total protein level for the two *FL* lines per unit of PEPC activity was much lower than the wild type. This could be an indication that the *KfCDF2_FL* lines have alterations in their leaf total protein content per unit of PEPC activity.

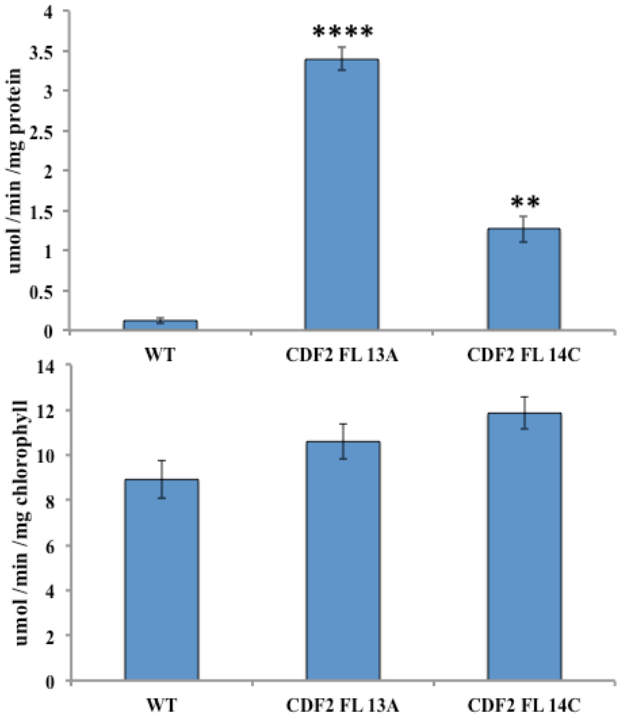

Figure 6.24. Activity level of PEPC during the dark period for WT and 2 *KfCDF2_FL* lines, **A.** Activity expressed per mg protein. **B.** Activity expressed relative to total chlorophyll content. Both *KfCDF2* lines showed vastly increased levels of PEPC activity compared to the wild type relative when expressed per mg of protein, but relative to total leaf chlorophyll content, neither line displayed a significant increase in activity. Two-way ANOVA Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars above one column represent significant differences to the corresponding WT value.

6.2.7.1.2 Measurement of malate sensitivity of PEPC

The apparent K_i of PEPC for malate was determined *in vitro* using enzyme-linked spectrophotometric assays with different levels of L-malate added to inhibit the extracted PEPC from the leaves of each line. These assay allowed the determination of the amount of L-malate required to achieve a 50 % inhibition of the control level of PEPC activity in the absence of L-malate, which is taken as a proxy for the apparent K_i of PEPC for malate inhibition. The malate sensitivity of PEPC is directly correlated with its phosphorylation state. More highly phosphorylated PEPC has a higher K_i for malate, around 3 mM in the dark in *K. fedtschenkoi* CAM leaves, and less strongly phosphorylated PEPC has a much lower K_i for malate, which in certain conditions can be lowere by about 10-fold (0.3 mM) in the light in *K. fedtschenkoi* CAM leaves (Nimmo *et al*., 1984). All three lines, wild type and the two *KfCDF2_FL* lines, showed this increase in K_i during the dark period (Table 6.5). Both transgenic lines showed a large decrease in K_i compared to the wild type at both 6 h light and dark using samples collected from a 12:12 LD cycle. For both lines, the light K_i was less than half of the value for the wild type, as was the dark K_i for *KfCDF2_FL_14C*. *KfCDF2_FL_13A* also showed a reduction in K_i of 1.3mM during the dark. This may explain why *KfCDF2_FL_14C* fixed less CO_2 during the dark period, but does not explain why *KfCDF2_FL_13A*, which had PEPC that was more sensitive to malate inhibition, managed to fix 35% more CO_2 during the dark period. The activity and regulation of RuBisCO in this line may therefore merit further investigation as it is the other major carboxylase enzyme in the leaves.

Table 6.4. Variation in the apparent K_i of PEPC for malate in rapidly desalted extracts of CAM leaves of WT and the *KFCDF2_FL* transgenic lines sampled in the middle of the 12 h light period (6 h L) and the middle of the 12 h dark period (6 h D).

Line	Light Malate K_i (mM)	Dark Malate K_i (mM)
WT	1.9	3.7
CDF2 FL 13A	0.7	2.4
CDF2 FL 14C	0.6	1.4

6.2.7.2 Enzyme assay results for *KfCDF2* RNAi knockdown lines

6.2.7.2.1 Total extractable activity of PEPC

Again, as with the over-expressing lines, both RNAi lines showed an increase in PEPC activity when expressed per mg protein, with *RNAi_19A* showing a 19.6-fold and *RNAi_26B* showing a 10.2-fold increase (Fig. 6. 25, 19A: $P < 0.0001$, 26B: $P = 0.0166$). This could be the basis for the ability of line *19A* fix much more CO_2 than the wild type, with a more rapid rate at the start of the dark period (Fig. 6.12). This highlights that immunoblotting needs to be conducted on this line to determine PEPC quantity and phosphorylation state of the PEPC throughout the dark period. Line *KfCDF2_RNAi_26B* fixed much less CO_2 during the dark period than the wild type (Fig. 6.13), and so the PEPC activity increase determined for this line was at odds with the CO_2 fixation data (Fig. 6.25).

When the total extractable PEPC activity was expressed per mg chlorophyll, line *KfCDF2_RNAi_26B* was found to have a decreased level of activity relative to the wild type (1.3-fold decrease), which correlated with the reduction in CO_2 fixation, and also by the immunoblot results for this line (Fig. 6.25). However, even when the activity was expressed per mg of chlorophyll, line *KfCDF2_RNAi_19A* still had significantly higher PEPC activity than the wild type with a 1.4-fold average increase (Fig. 6.25, $P = 0.0413$). This suggested that line *19A* had higher PEPC activity than

the wild type, which correlated with the corresponding CO_2 exchange data which suggested that this line was fixing more CO_2 in the dark period (Table 6.3).

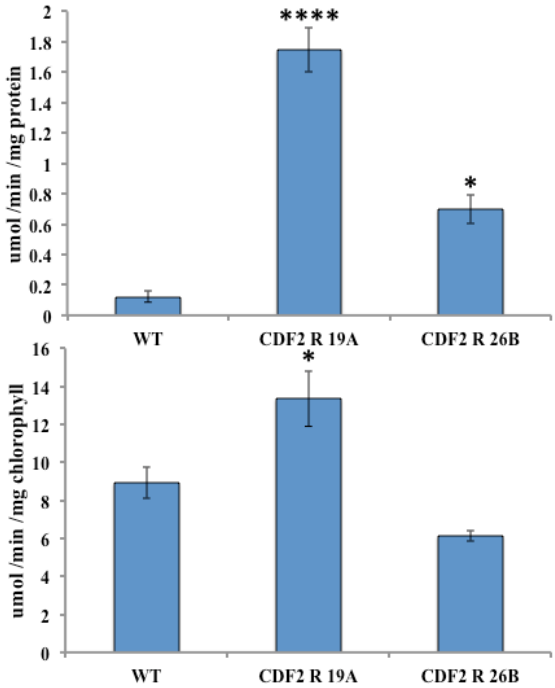

Figure 6.25. Total extractable activity of PEPC during the dark period for WT and 2 *KfCDF2_RNAi* lines, **A.** Activity expressed per mg total leaf protein. **B.** Activity expressed relative to total leaf chlorophyll content. Both *KfCDF2_RNAi* lines show vastly increased levels in activity compared to the wild type when values were normalised per mg of protein, but when expressed relative to the leaf chlorophyll content, only *KfCDF2_RNAi_19A* had an increase in activity. Two-way ANOVA Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars above one column represent significant differences to the corresponding WT value

6.2.7.1.2 Measurement of malate sensitivity of PEPC

Both transgenic lines showed a decrease in the apparent K_i of PEPC for malate relative to the wild type, but *KfCDF2_RNAi_19A* showed a much smaller difference compared to wild type (Table 6.6). *KfCDF2_RNAi_26B* shows a larger drop in K_i compared to wild type. During the light, PEPC in line *26B* was 50 % inhibited by less than half the malate concentration required to inhibit 50 % of the PEPC activity in the

wild type (Table 6.6), and during the dark the PEPC extracted from this line was also inhibited by 50 % by a lower concentration of malate (Table 6.5).

These results suggested that the PEPC in both lines would be inhibited by lower concentrations of malate than were required to inhibit the enzyme in the wild type, suggesting there had been a decrease in the phosphorylation of PEPC in both lines.

Table 6.5. Variation in the apparent K_i of PEPC for malate in rapidly desalted extracts of CAM leaves of WT and the *KfCDF2_RNAi* transgenic lines sampled in the middle of the 12 h light period (6 h L) and the middle of the 12 h dark period (6 h D).

Line	Light Malate K_i (mM)	Dark Malate K_i (mM)
WT	1.9	3.7
CDF2 RNAi 19A	1.3	3.1
CDF2 RNAi 26B	0.8	2.4

6.2.8 The impact of drought stress on the growth and stress-related physiology of the *KfCDF2* transgenic lines

CAM is widely argued to play an important role in increasing plant water use efficiency and productivity in water limited environments. To investigate the performance of the *KfCDF2* transgenic lines under drought-stress conditions, where efficient and optimised CAM would be expected to become more important to the growth and survival of the plants, large, developmentally synchronised clonal populations of the *KfCDF2* transgenic lines and wild type were grown from leaf margin plantlets in 16:8 greenhouse conditions. All plants were grown side-by-side under well-watered conditions until they were 6-months-old, at which point the population of each line was separated in half. Half of the plants for each line were maintained under well-watered conditions (5 mins of watering daily by automoatic flooding of the bench), whilst the other half of the plants received no further water.

After 31 days in the two conditions, the plants were harvested and various measurements were recorded.

6.2.8.1 Drought effects on *KfCDF2* full-length overexpressing lines

KfCDF2_FL_13A had 7 plants in each conditions and *KfCDF2_FL_14C* had 9 plants. Numbers varied because some lines grew better than others and only clones with uniform growth were selected for the experiment.

6.2.8.1.1 Above-ground biomass

For all lines in the FL over-expressers growth experiment, there were significant differences between fresh weights relative to the wild type (Fig. 6.26A). *KfCDF2_FL_13A* had a significant reduction in the fresh weight of its above ground biomass in well-watered conditions, whereas line *KfCDF2_FL_14C* showed a slight increase in weight, although this was not statistically significant (Well-watered: WT = 371 g, *13A* = 244 g, *14C* = 422 g; Drought: WT = 75.6 g, *13A* = 80.0 g, *14C* = 195 g). However, the dry weight data revealed a significant difference for *FL_14C,* which had an increase in shoot dry weight relative to the wild type under the drought conditions (Fig. 6.26C; Well-watered: WT = 18.1 g, *13A* = 5.8 g, *14C* = 15.0 g; Drought: WT = 10.0 g, *13A* = 3.6 g, *14C* = 17.7 g). Interestingly, both lines also showed a significant increase in % moisture in the plants under the drought conditions despite soil water content being approximately equal in all lines (Appendix: Fig. S6.1 & Table S6.1) (Fig. 6.26E; Well-watered: WT = 95.1 %, *13A* = 97.6 %, *14C* = 96.5 %; Drought: WT = 86.8 %, *13A* = 95.5 %, *14C* = 90.0 %). This suggested that the over-expression of *KfCDF2* in some way led to an increase in WUE for these plants,

especially as *KfCDF2_FL_14C* was able to also grow significantly more than WT during the drought-stress treatment.

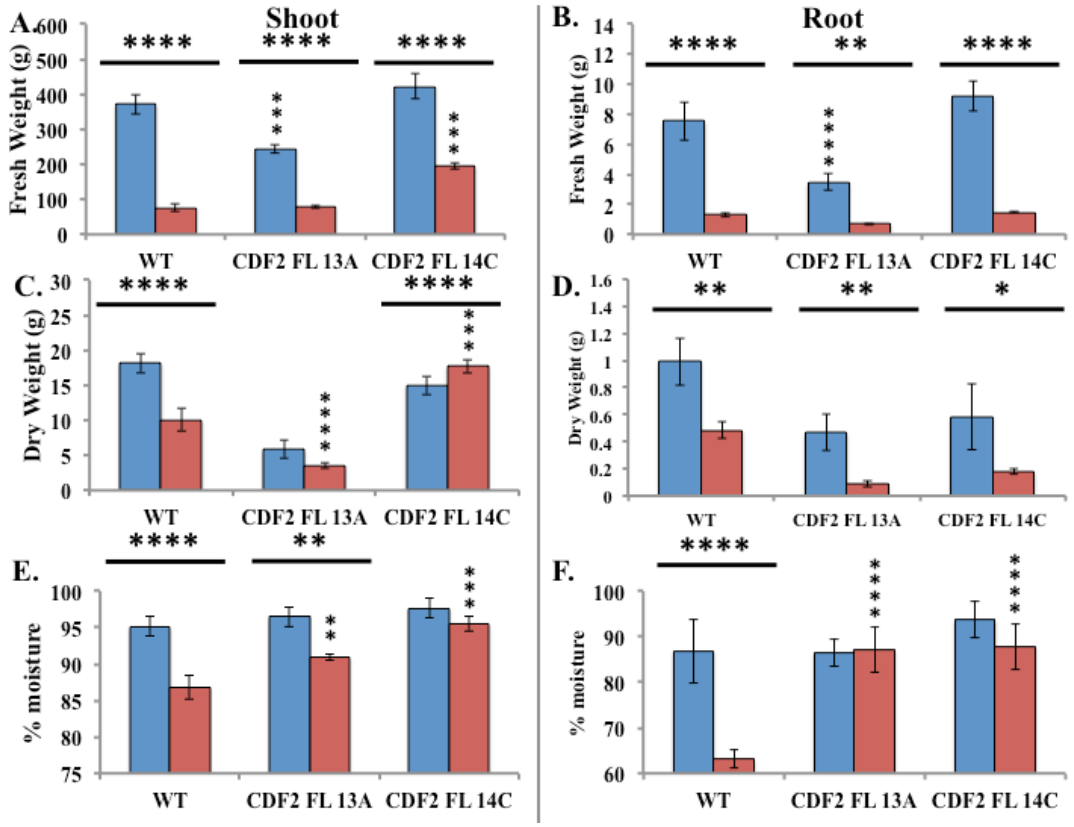

Figure 6.26. Fresh and dry weight and relative moisture content for well-watered and drought-stressed WT and FL lines. Blue bars show well-watered lines, and red bars show drought-stressed lines. A, C & E show data collected for above ground weight, and B, D & F show data for below ground tissues (roots). A and B show fresh weight of samples at time of harvest, C and D, show dry weight after tissues were placed in a drying oven at 70˚C until a constant weight was reached. E and F show the % moisture held within the plant; calculated from the fresh and dry weights. A Two-Way ANOVA Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one column, e.g. well-watered/ drought represent significant differences compared to the corresponding WT value.

6.2.8.1.2 Below-ground root biomass

For both *FL* lines there were significant difference between fresh weights for the roots (Fig. 6.26B). When comparing to wild type lines, *KfCDF2_FL_13A* had significantly reduced fresh weight in well-watered conditions for roots, whereas *14C* had a slightly increased root fresh weight (Well-watered: WT = 7.5 g, *13A* = 3.5 g, *14C* = 9.2 g; Drought: WT = 1.3 g, *13A* = 0.7 g, *14C* = 1.48 g). From the root dry weight data, in both well-watered and drought-stress conditions, both *FL* lines produced less root material than the wild type, although none of these reductions were significant (Fig. 6.26D; Well-watered: WT = 0.99 g, *13A* = 0.47 g, *14C* = 0.58 g; Drought: WT = 0.48 g, *13A* = 0.08 g, *14C* = 0.62 g). When percentage moisture was calculated for the roots, only the wild type showed a significant difference between drought-stress and well-watered roots (Fig. 6.26F). The other two lines showed greatly increased moisture levels in drought-stressed roots relative to the wild type (Well-watered: WT = 86.8 %, *13A* = 86.5 %, *14C* = 93.7 %; Drought: WT = 63.1 %, *13A* = 87.2 %, *14C* = 86.9 %). The increases in *KfCDF2* expression in these transgenic *K. fedtschenkoi* lines thus led to an increase in the whole plant's ability to maintain its moisture content during drought.

6.2.8.1.3 Induction of UV protective anthocyanin pigments and other pigmentation responses to drought-stress

As wild type *K. fedtschenkoi* plants progress into a drought-stressed state, they accumulate anthocyanin pigments in their leaves. These pigments are believed to help prevent UV damage, and the pathway of anthocyanin biosynthesis has been shown to be regulated by the circadian clock in Arabidopsis. The transgenic lines studied here were all manipulated for CAM-induced TFs whose transcript abundance oscillated

with a robust circadian rhythm. It was therefore important to investigate the ability of the *KfCDF2_FL* lines to produce anthocyanins in response to drought stress. Both over-expresser lines displayed significant differences in the level of total leaf anthocyanins between well-watered and drought-stress conditions. In each condition though, all three lines produced roughly the same amount of leaf anthocyanins in each condition. Line *13A* showed a small decrease, whereas line *14C* had a small increase in anthocyanins levels produced in well-watered and drought-stressed conditions (Fig 6.27; Well watered: WT = 0.0013 µmol cm^{-2}, *13A* = 0.0016 µmol cm^{-2}, *14C* = 0.0007 µmol cm^{-2}, Drought: WT = 0.0083 µmol cm^{-2}, *13A* = 0.0094 µmol cm^{-2}, *14C* = 0.0077 µmol cm^{-2}).

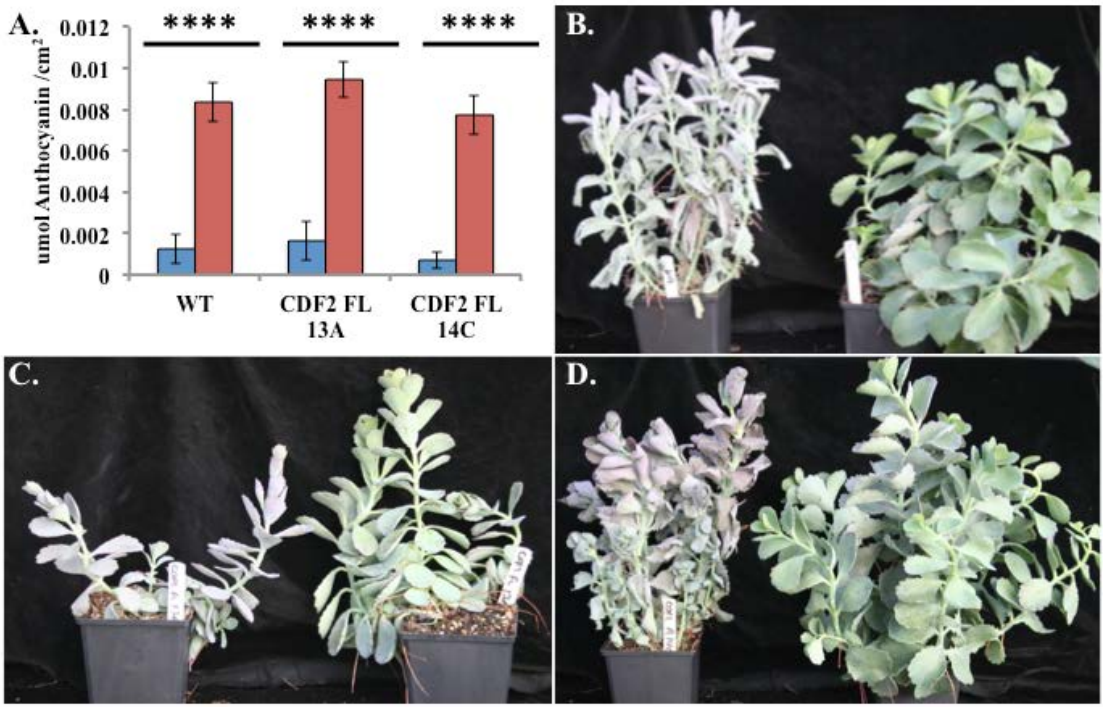

Figure 6.27. Variation in the level of anthocyanin leaf pigments in response to drought-stress for WT and *KfCDF2_FL_13A* and *14C*. **A.** Quantified levels of Anthocyanins in WT and the *KfCDF2 FL* lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one column, e.g. drought, represent significant differences to the corresponding WT value. **B.** Images of drought-stressed (Left) and well-watered (Right) WT. **C.** Images of drought-stressed (Left) and well-watered (Right) *KfCDF2_FL_13A*. **D.** Images of drought-stressed (Left) and well-watered (Right) *KfCDF2_FL_14C*.

Other leaf pigments associated with photosynthetic light harvesting were also measured, namely chlorophyll A, B and carotenoids. All three pigments increased significantly in drought-stressed leaves, even after leaf area changes associated with drought had been accounted for (Fig 6.28A). Despite many of the results not showing significant differences, there was still a trend whereby in *KfCDF2_FL* lines, in both drought and well-watered conditions, there were lower amounts of each pigment. In well-watered conditions, there were significant reductions in Chlorophyll A measurements in both *FL* lines (Fig. 6.28; *13A*: P = 0.0032, *14C*: P = 0.0343). *KfCDF2_FL_13A* also showed a significant decrease in Chlorophyll B and Carotenoids under drought (Fig. 6.28B; ChlB = P = 0.0070, Carotenoids: P <0.0001), resulting in there not being significant differences between the level of Chlorophyll B or carotenoids for this line when comparing well-watered and drought-stressed conditions (Fig. 6.28C).

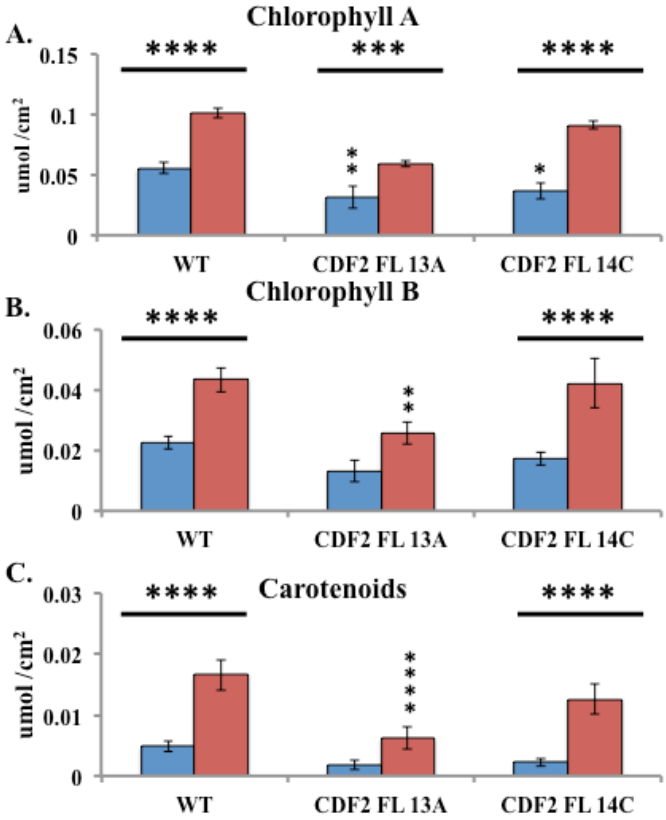

Figure 6.28. Variation in the levels of photosynthetic pigment in response to drought-stress in WT and *KfCDF2_FL* lines. **A.** Changes in the levels of Chlorophyll A in response to drought. **B.** Changes in the levels of Chlorophyll B in response to drought. **C.** Changes in the levels of Carotenoids in response to drought. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; **** shows P = ≤ 0.0001. Stars over bars at the top of the graph represent lines that are significantly different between treatments.

6.2.8.1.4 Metabolite levels in drought conditions

It was also important to assess the impact of drought-stress on the daily CAM cycle. In well-watered conditions, all lines turned over malate efficiently between the end of the dark and the end of the light, with the *FL* lines accumulating less malate than the wild type (Fig 6.29A). However, whilst in the previous malate measurements (Fig. 6.14), the FL lines accumulated significantly less malate than the wild type by the end of the dark period, here the equivalent well watered plants of the *FL* lines did not

show a significant decrease in the amount of malate accumulated by dawn relative to the wild type.

After drought stress, the wild type accumulated similar amounts of malate at the end of the dark period to the levels it accumulated under in well-watered conditions, whereas both over-expresser lines showed large and significant reductions in the amount of malate they accumulated by dawn (Fig. 6.29D). This has resulted in vast differences between the lines Δ-malate values, with wild type turning over only slightly more malate than *KfCDF2_FL*s in well watered conditions (WT = 2.4 μmol gFW^{-1}, 13A & 14C = 2.1 μmol gFW^{-1}), whilst during drought wild type turns over a great deal more (WT = 90.5 μmol gFW^{-1}, 13A = 3.9 μmol gFW^{-1}, 14C = 45.3 μmol gFW^{-1}). This therefore suggested that both *FL* lines may have transitioned into performing a degree of CAM-idling, where stomata are maintained closed throughout the light/ dark cycle, and the only nocturnal CO_2 available for PEPC fixation comes from respired CO_2. This possibility would need to be tested further by making measurements on the stomatal conductance of the drought-stressed FL lines. However, the LICOR 6400XT IRGA was not available during this experiment, but the experiment could now be repeated and stomatal measurements made using the LICOR IRGA system.

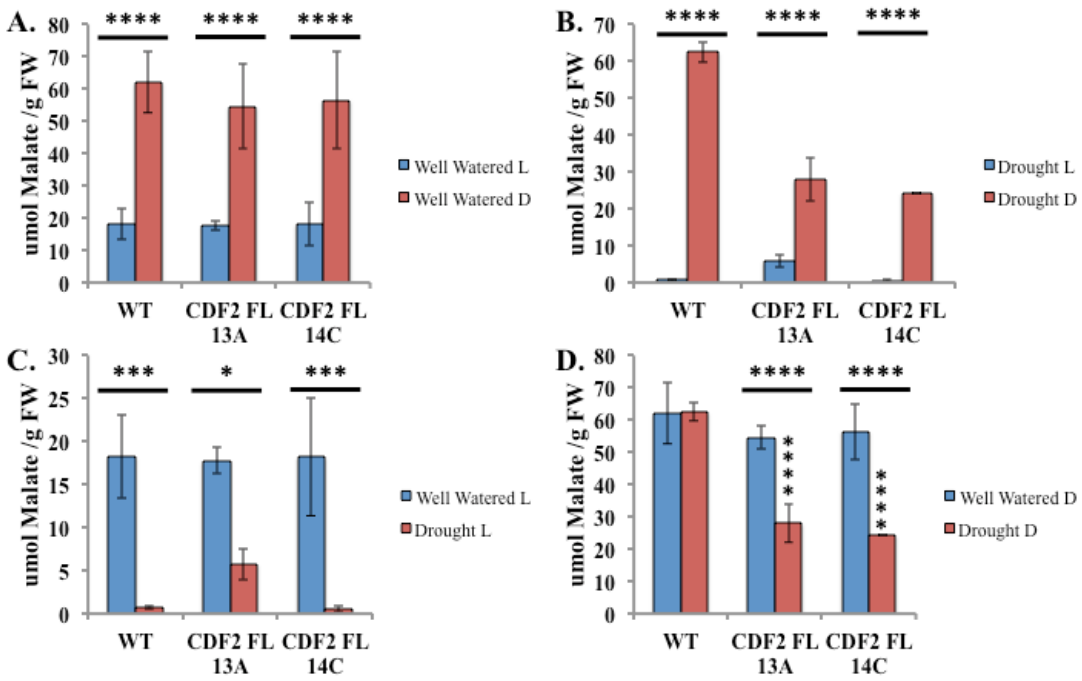

Figure 6.29. Daily fluctuations in malate content of CAM leaves of the WT and *FL* over-expressor lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. well-watered and Drought, and stars above one column, e.g. drought, represent significant differences from the corresponding WT value.

6.2.8.2 Drought effects on *KfCDF2* RNAi knockdown lines

Measurements were made using 8 plants of line *KfCDF2_RNAi_19A* under both the well-watered and drought-stressed conditions, and 9 plants of line *KfCDF2_RNAi_26B* were grown under both conditions.

6.2.8.2.1 Above-ground biomass

All tissues were harvested after 31-days in either well-watered or drought-stressed conditions. The two RNAi lines and wild type all showed significant differences between well-watered and drought-stress conditions for their above ground fresh weights (Fig. 6.30A; Well watered: WT = 371 g, *19A* =326 g, *26B* = 396 g; Drought: WT = 75.8 g, *19A* = 77.8 g, *26B* = 79.7 g), dry weights (Fig. 6.30C; Well watered:

WT = 18.1 g, *19A* = 16.5 g, *26B* = 14.6 g; Drought: WT = 1.00 g, *19A* = 9.1 g, *26B* = 9.7 g), and percentage moisture (Fig. 6.30E; Well watered: WT = 95.1 %, *19A* = 95.8 %, *26B* = 95.5 %; Drought: WT = 86.8 %, *19A* = 88.3 %, *26B* = 87.8 %). However, compared to the wild type, neither line showed any significant differences for any of the measured parameters for the above ground tissues. It is however important to note that, compared to wild type in drought, both lines showed slight increase in their percentage moisture in their above ground tissues (Fig. 6. 30E).

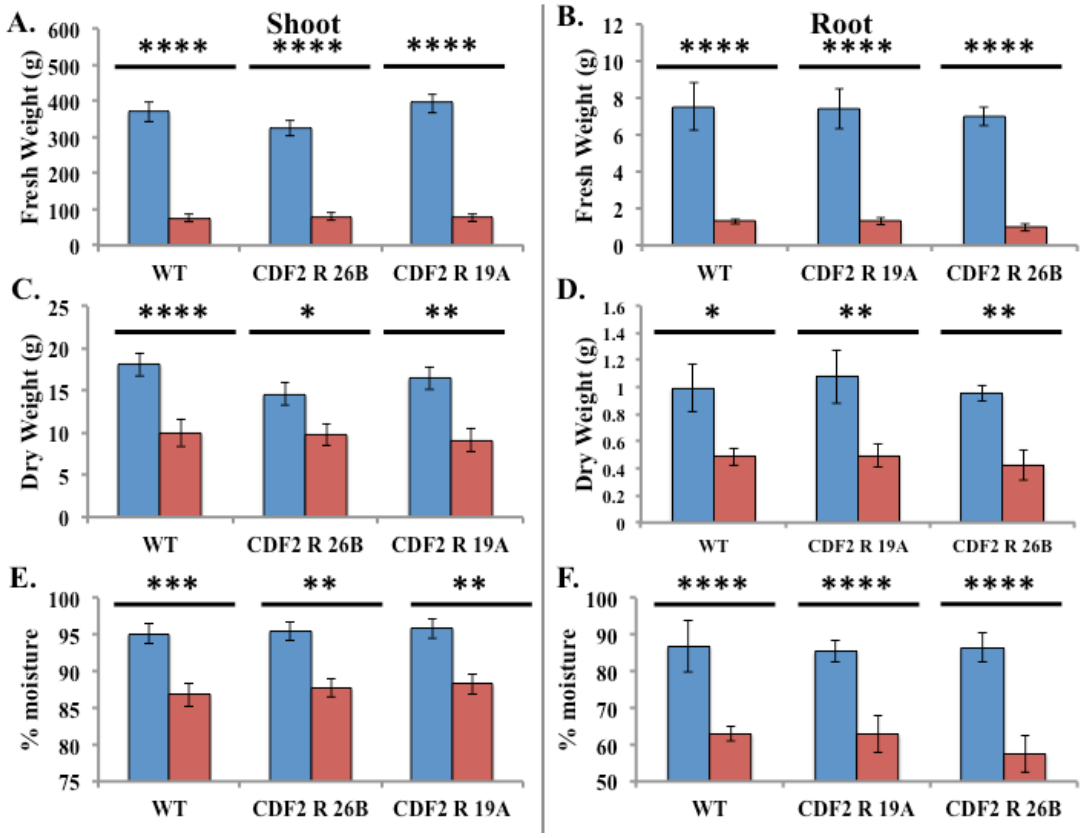

Figure 6.30. Fresh and dry weight and relative moisture content for well-watered and drought-stressed WT and RNAi lines. Blue bars show well watered lines, and red bars show drought-stressed lines. A, C & E show data collected for above ground weight, and B, D & F show data for below ground tissues (roots). A and B show fresh weight of samples at time of harvest, C and D, show dry weight after tissues were placed in a drying oven at 70˚C until a constant weight was reached. E and F show the % moisture held within the plant; calculated from the fresh and dry weights. A Two-Way ANOVA Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different eg. L and D, and stars above one column, e.g. well-watered/ drought represent significant differences compared to the corresponding WT value.

6.2.8.2.2 Below-ground root biomass

Roots were also removed from each harvested plant and weighed when both fresh and dry. For each line there were again significant differences in weight between well-watered and drought-stressed conditions for fresh weight (Fig 6.30B; Well-watered: WT = 7.56 g, *19A* = 7.45 g, *26B* = 7.00 g; Drought: WT = 1.31 g, *19A* = 1.32 g, *26B* = 1.00 g), Dry weight (Fig 6.30D; Well-watered: WT = 1.00 g, *19A* = 1.08 g, *26B* = 0.95 g; Drought: WT = 0.49 g, *19A* = 0.49 g, *26B* = 0.42 g), and percentage moisture (Fig. 6.30F; Well-watered: WT = 86.8 %, *19A* = 85.5 %, *26B* = 86.4 %; Drought: WT = 63.1 %, *19A* = 62.8 %, 26B = *57.5* %). Despite the fact that there were no significant differences between the RNAi lines and the wild type, *KfCDF2_RNAi_26B* did show a slight decrease in weight and also percentage moisture in roots compared to the wild type (Fig. 6.30F).

6.2.8.2.3 Induction of UV protective anthocyanin pigments and other pigmentation responses to drought-stress

In well watered conditions both RNAi lines showed lower amounts of total leaf anthocyanins than the wild type, although these differences were not found to be significant (WT = 0.0013 μM cm^{-2} , *19A* = 0.0006 μM cm^{-2} , *26B* = < 0.0001 μM cm^{-2}). Under the drought conditions, both *KfCDF2_RNAi* lines showed significantly lower amounts of anthocyanin being produced (both *19A* and 26B compared to WT: $P < 0.0001$). These figures were supported visually by the photographs presented in Figure 6.31 B, C and D. In Figure 6.31B, the drought-stressed wild type (Left) has a purple tinge, whereas in Figures 6.31C and D, the *KfCDF2 RNAi* lines under drought-stress had not developed such as strong purple tinge to their leaf colour.

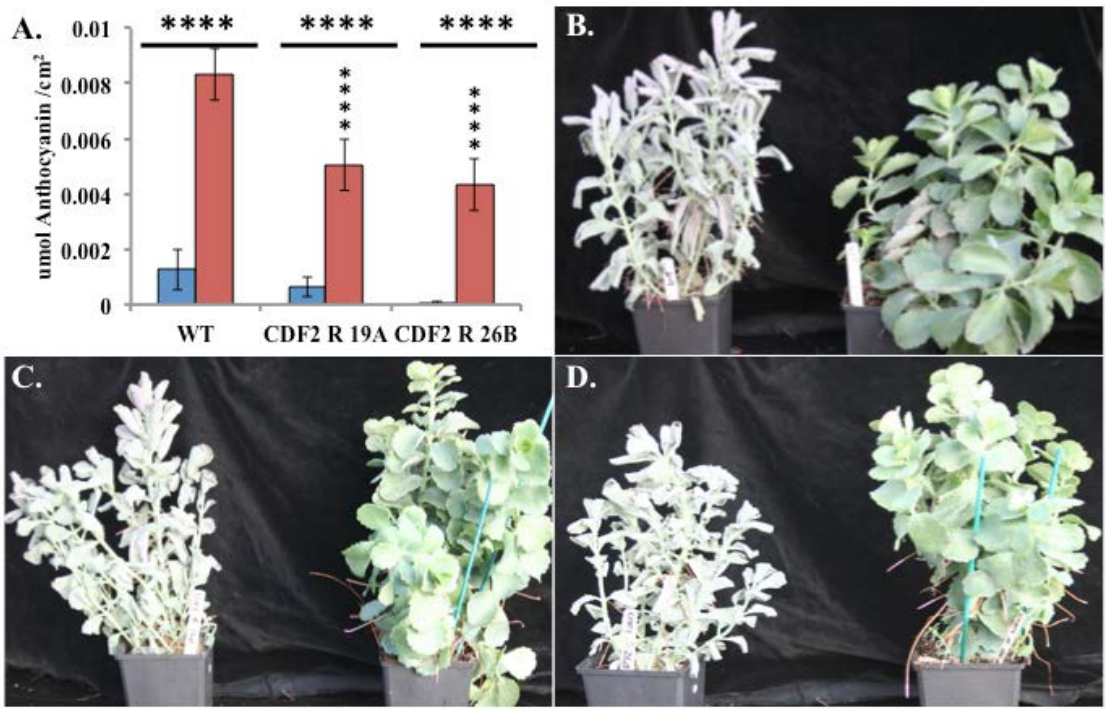

Figure 6.31. Variation in the level of anthocyanin leaf pigments in response to drought-stress for WT and *KfCDF2_RNAi_19A* and *26B*. **A.** Quantified levels of Anthocyanins in WT and the *KfCDF2_RNAi* lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. L and D, and stars above one column, e.g. drought, represent significant differences to the corresponding WT value. **B.** Images of drought-stressed (Left) and well-watered (Right) WT. **C.** Images of drought-stressed (Left) and well-watered (Right) *KfCDF2_RNAi_19A*. **D.** Images of drought-stressed (Left) and well-watered (Right) *KfCDF2_RNAi_26B*.

Photosynthetic pigments were also measured for the well-watered and drought-stressed plant from this experiment (Fig. 6.32). Despite a significant increase in all three pigments comparing well-watered to drought-stressed conditions, there were no differences between the two RNAi lines and the wild type (Fig. 6.32). This suggests that in these *KfCDF2* RNAi lines did not have any major changes in their photosynthetic light harvesting capacity.

These results for the leaf pigment levels in the RNAi lines were the exact opposite of the results for the *FL* lines, as the *KfCDF2_FL* lines showed no significant differences

in anthocyanin, but significant differences from the wild type for the photosynthetic pigments.

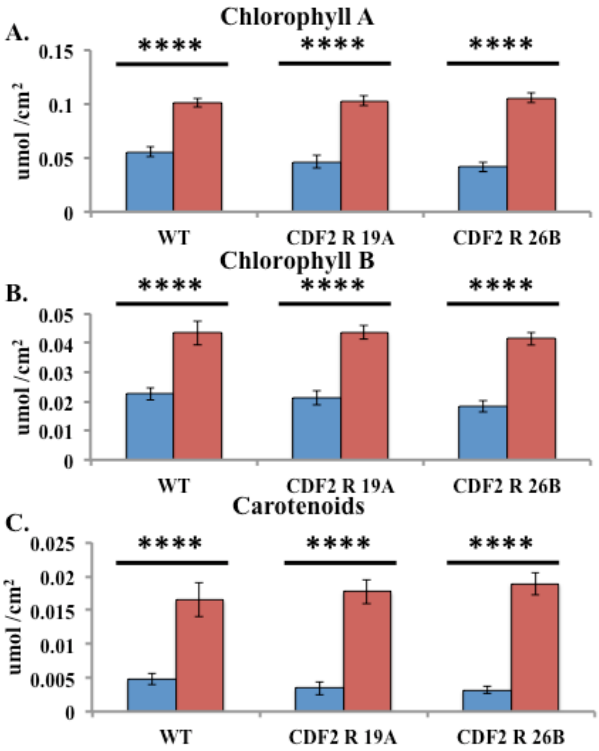

Figure 6.32. Variation in the levels of photosynthetic pigment in response to drought-stress in WT and *KfCDF2_RNAi* lines. **A.** Changes in the levels of Chlorophyll A in response to drought. **B.** Changes in the levels of Chlorophyll B in response to drought. **C.** Changes in the levels of Carotenoids in response to drought. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; **** shows P = ≤ 0.0001. Stars over bars at the top of the graph represent lines that are significantly different between treatments.

6.2.8.2.4 Metabolite levels in drought conditions

The level of malate was assayed to determine the effect of the drought-stress treatment on CAM. In well-watered conditions, there was a significant turnover of malate over the 24 h cycle in all lines (Fig. 6.33A; P <0.0001). Under the drought-stressed conditions, there was a decrease in both light and dark accumulation of malate for the RNAi lines (Fig. 6.33B & C). The dark amount of malate accumulated was significantly decreased in both RNAi lines relative to the wild type following the drought-stress treatment (Fig. 6.33D; P <0.0001). Drought conditions have therefore

affected malate turnover, with wild type in well watered conditions turning over slightly more malate during the 24h cycle (WT = 43.9 µmol gFW^{-1}, 19A = 36.4 µmol gFW^{-1}, 26B = 43.4 µmol gFW^{-1}), whereas in drought wild type now cycles much more malate each day (WT = 61.8 µmol gFW^{-1}, 19A = 31.6 µmol gFW^{-1}, 26B = 22.03 µmol gFW^{-1}). This indicates atleast phase IV has been abolished for wild type leading to the line having to decrboxylate nearly all its available malate. This may also indicate that these transgenic lines have entered CAM-idling, whereby they were only able to re-fix respired CO_2 behind stomata that remained closed through the light/ dark cycle.

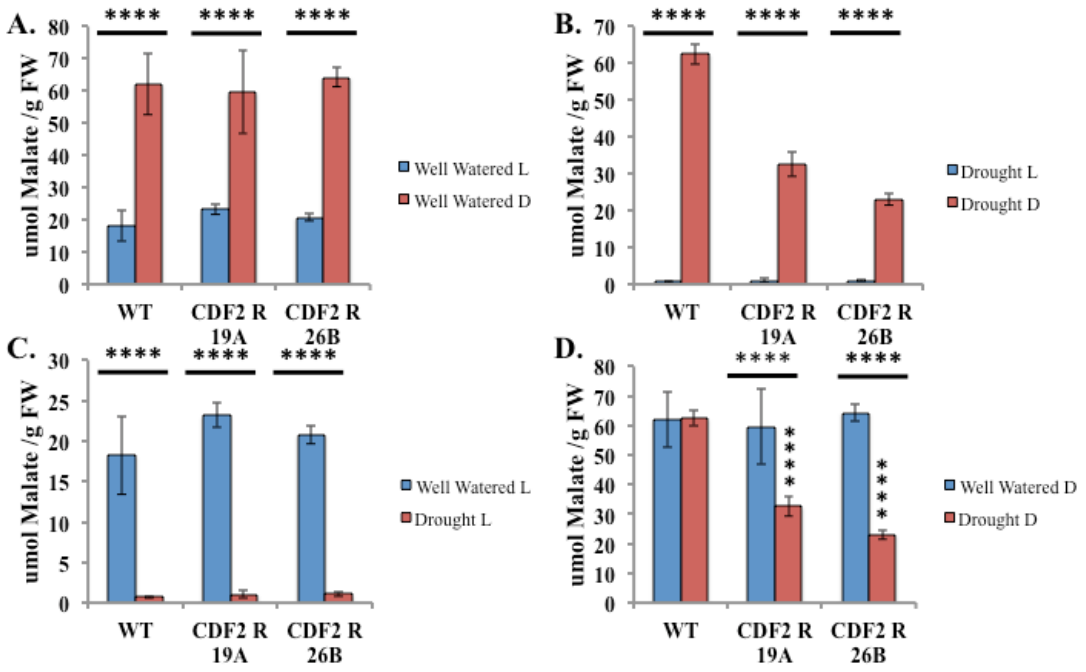

Figure 6.33. Daily fluctuations in malate content of CAM leaves of the WT and *KfCDF2_RNAi* knockdown lines. Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars over bars at the top of the graph represent groups that are significantly different e.g. well-watered and Drought, and stars above one column, e.g. drought, represent significant differences from the corresponding WT value.

6.2.9 Succulence in CDF2 transgenic lines

From the drought-stress experiment, it was clear that line *KfCDF2_FL_13A* did not lose water as rapidly as other lines, suggesting this line was able to maintain its leaf succulence for longer under drought-stress conditions compared to the wild type. This may in turn suggest that this line could survive under drought for a longer period than the wild type.

6.2.9.1 Succulence in well-watered conditions

Each *KfCDF2* transgenic line and the wild type were grown up from developmentally synchronised leaf-margin plantlets under 16:8 greenhouse conditions for 6 months. Succulence was measured at day 0, when all plants were well watered due to a daily watering regime. Watering continued throughout the 90 day sampling period, and all above-ground plant material was harvested for succulence measurements (g of water per cm^2) on days 14, 28, 42 and 90.

The *KfCDF2* transgenic lines displayed a gradual increase in leaf succulence throughout the 90 days of this experiment, whereas the wild type did not maintain the same rate of increase (Fig. 6.34). This suggested that the perturbation of the expression level of *KfCDF2* in these transgenic lines, both the increased expression in the *FL* lines, and the decreased level of expression in the RNAi lines, altered the regulation of water storage in the leaves during their development under well-watered conditions in the greenhouse. It is particularly noteworthy that the strongest over-expresser line, *FL_13A*, showed the highest value for relative leaf succulence at the end of this 90 day experiment.

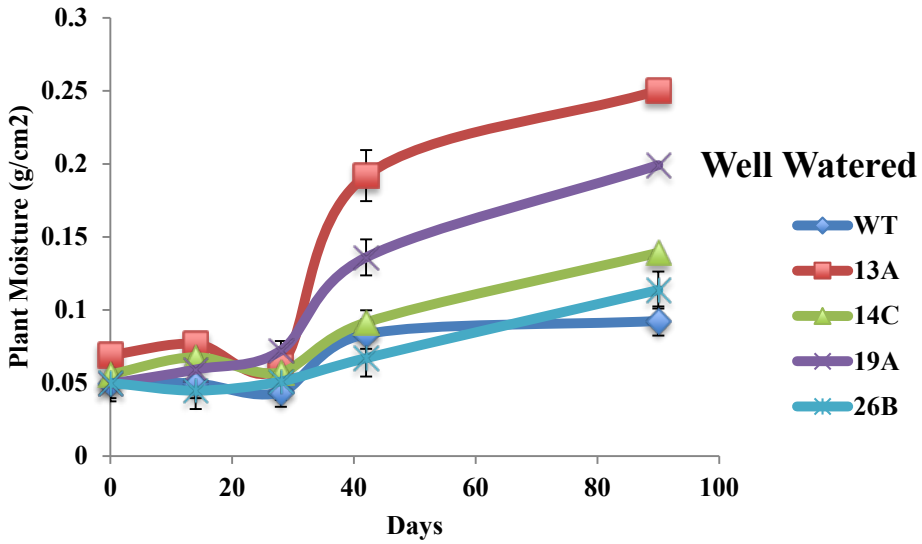

Figure 6.34. Succulence calculated for 6-month-old *KfCDF2* transgenic lines and wild type over 90 days of growth in well-watered, greenhouse conditions (16:8 LD). The four transgenic lines all increased their leaf succulence at a faster rate than the wild type. All lines increased their leaf succulence over the 90 days. *KfCDF2_FL_13A* showed the greatest increase in its leaf water content relative to its leaf area.

6.2.9.2 Leaf succulence changes in the wild type and *KfCDF2* transgenic lines under long-term drought treatment

During 90 days drought-stress without watering, line *KfCDF2_FL_13A* maintained a slightly higher level of succulence (g water per cm^2) than the wild type (Fig. 6.35). The other *KfCDF2* lines also maintained slightly higher leaf succulence than the wild type, especially after 14 and 28 days of water with-holding (Fig. 6.35). Leaves of *KfCDF2_FL_13A* were more succulent than the wild type at the beginning of the experiment, and this line was able to maintain this advantage throughout the 90 day period without water, but by the 90 day point, the succulence advantage of line *FL_13A* was greatly diminished relative to the start of the experiment between days 14 and 28 (Fig. 6.35).

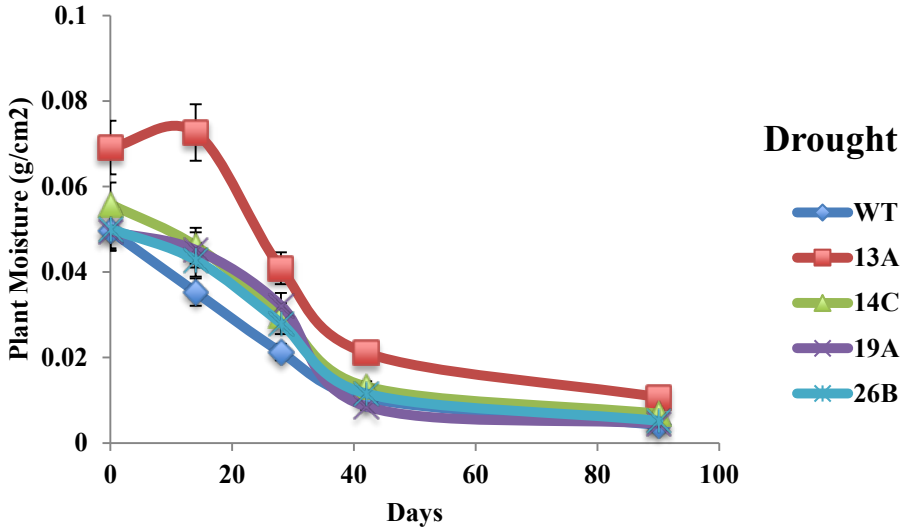

Figure 6.35. Succulence calculated for 6-month-old *KfCDF2* transgenic lines and wild type over 90 days of growth without watering under green house conditions (16:8 light/ dark). *KfCDF2_FL_13A* showed higher leaf succulence than the other lines throughout the 90 days. All transgenic lines performed better than the wild type during the early part of the drought treatment, and maintained increased succulence for the first 28 days with no water.

In both well-watered and drought-stressed conditions, the wild type plants were less succulent than the transgenics both at the start of the experiment and after 90 days, which did support the idea that changes in *KfCDF2* expression led to increased leaf succulence. This could be linked to the fact that the *KfCDF2* lines did not produce as much vegetative growth as the wild type lines, especially line *FL_13A* (Fig. 6.36), and its possible that this may be linked to the provisioning of resources towards the constitutive long-day flowering that occurred in these plants.

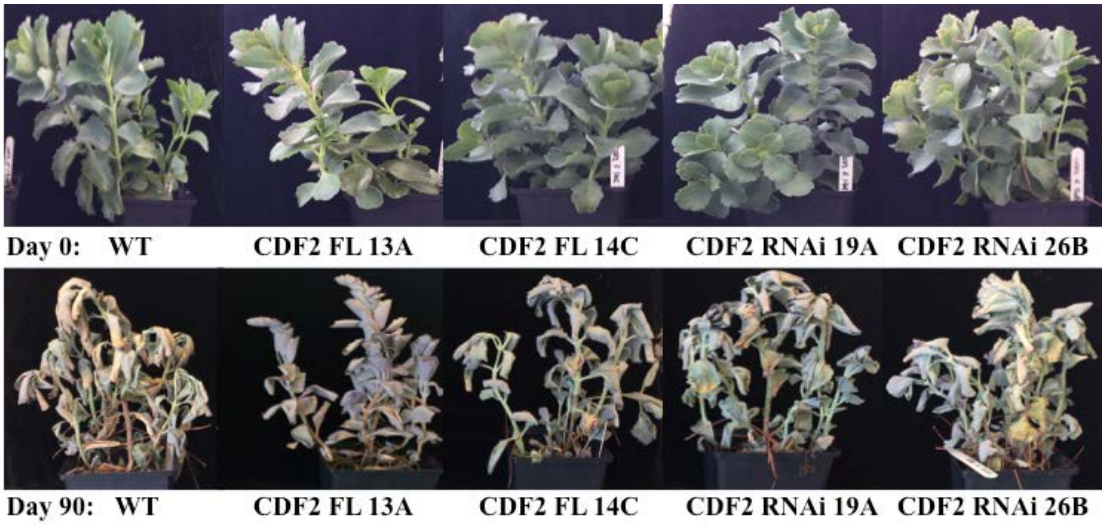

Figure 6.36. WT and *KfCDF2* transgenic lines on day 0 and day 90 of the drought time course. All lines look approximately the same on day 0, except *KfCDF2_FL_13A*, which had less vegetative growth. By day 90, the effects of drought were more severe for the wild type than for the other lines. Line *KfCDF2_FL_13A* was attempting to flower (note the bending over of the stem at the growing point) at the 90 day time point, despite the severe drought-stress conditions it was enduring. This indicates that line *FL_13A* was still able to grow and develop at a time when the wild type was close to complete senescence.

From this drought experiment, we can compare plant appearance between lines. Wild type (WT) suffered most under this prolonged drought stress treatment, as the plants had many more brown and dead leaves, and the stems were also bending more (a sign of severe wilting in *K. fedtschenkoi*) (Fig. 6.36). This supports the data for the much lower levels of water in the leaves of the wild type plants compared to the transgenic *KfCDF2* lines (Fig. 6.35).

All lines, apart from *KfCDF2_FL_13A,* appear to have greatly reduced purple colouring (due to anthocyanins) after 90 days of drought, despite being under increasing amounts of stress (Fig. 6.36). This may signify that the anthocyanins had been turned over as part of the plant's last-ditch efforts to stay alive during such a long-term drought treatment. When the plants are this drought-stressed, it may be that non-essential pathways and processes are stopped, and anthocyanins may be re-assimilated as part of the battle to keep cellular machinery alive.

KfCDF2_FL_13A was much less strongly affected by drought compared to the other lines and was still able to maintain a high leaf anthocyanin content as judged by the purple colouration of the leaves (Fig. 6.36). *KfCDF2_FL_13A* was developing its shoot apical meristem towards flowering, as can be seen by the hooked stem (Fig. 6.36). Closer inspection of the *KfCDF2_FL_13A* plants did reveal that the older leaves near the bottom of the stem had started to senesce and die, whilst newer leaves near the top of the stem (LP1 – 7), remained turgid (Fig. 6.36). This indicated that the remaining water resources were focussed in the younger more recently developed leaves and around the meristem, possibly to enable the developmental progression towards flowering to continue.

Finally, after the 90 days of drought, the remaining plants were watered. Photographs were then taken 24, 48 and 168 hours (1 week) after watering restarted to look at recovery (Fig. 6.37). The *KfCDF2_FL* lines both showed fast recovery after 24 h, whereas *KfCDF2_RNAis 19A* and *26B* did not appear to recover much at all, and shed many leaves. The wild type took longer to recover than the FL lines too, but did not shed as many leaves as *RNAi_26B*.

After 1 week of watering, both FL lines had fully recovered from drought (at least visually; Fig. 6.37), whereas the RNAi lines and wild type had not. The wild type and RNAi lines only had succulent leaves close to the meristem (Fig. 6.37). The other older leaves further down the stem were either senescent or had already been shed from the plant and died (Fig. 6.37).

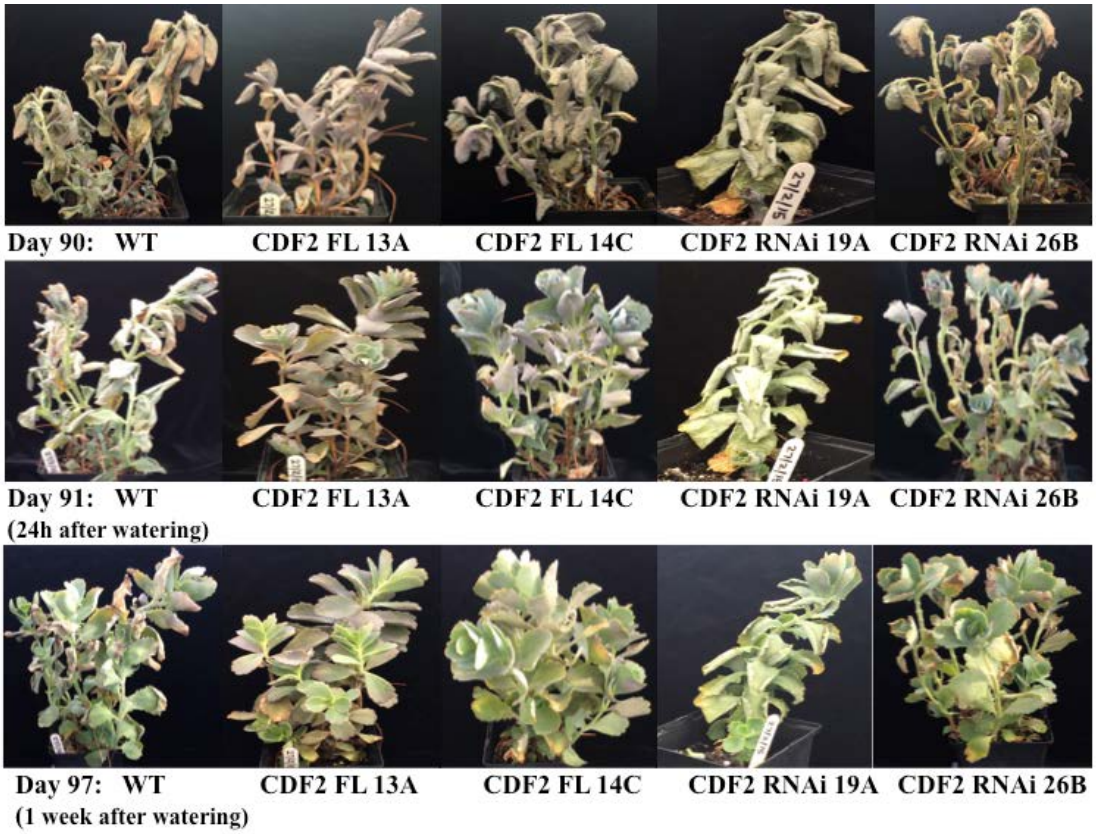

Figure 6.37. WT and *KfCDF2* transgenic lines on day 90 of the drought time course (before watering), day 91 (24 h after re-watering) and day 97 (1 week after re-watering). Both *KfCDF2_FL* lines recovered much more quickly than the wild type, whereas both RNAi lines did not recover as quickly or comprehensively, except for their youngest leaves at the top of the stem. Line *RNAi_26B* lost many leaves after watering.

This phenotype of *KfCDF2* over-expression conferring increased drought tolerance on the plants was further investigated using the confirmed HA-tagged *KlCDF2* over-expresser lines of *Kalanchoë laxiflora* that were generated in the last 12 months of this project in the hope that some ChIP-Seq work could be undertaken to identify the genomic targets of each TF. *K.laxiflora* wild type and *HA-KlCDF2_FL* over-expresser lines were subjected to drought-stress for approximately three weeks. These *HA-KlCDF2_FL* over-expresser lines showed a more marked flowering phenotype than the *FL* lines of *K. fedtschenkoi*, but also displayed greater drought tolerance, and showed a much higher productivity than the wild type (Fig. 6.38). This supports the

proposal that function of *KfCDF2* is conserved in *K. laxiflora*, which is a very closely related species (Fig.6.38).

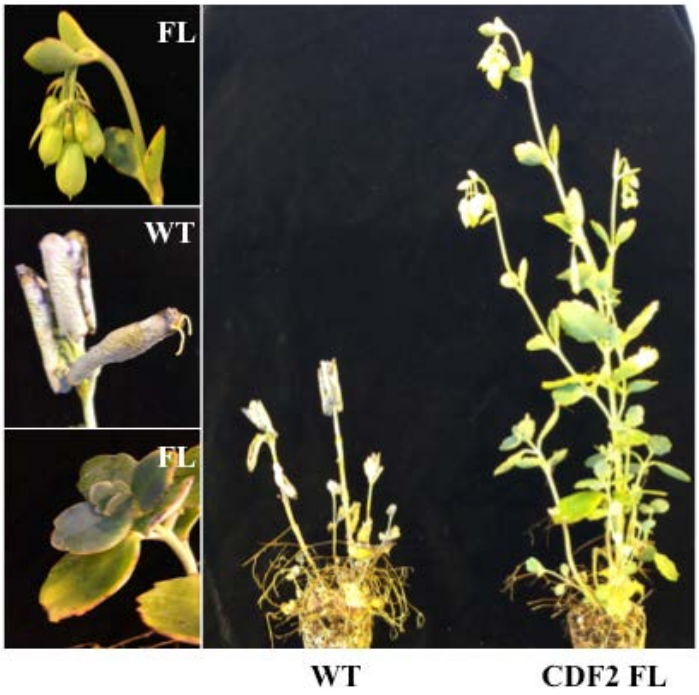

Figure 6.38. *K. laxiflora* wild type (WT) and *KfCDF2_FL* overexpressing line (*CDF2_FL*) after approximately three weeks of drought. The wild type plant had shed most of their leaves, and they were purple and had started to produce plantlets from their leaf margins. By contrast, the *KfCDF2_FL* line still had turgid leaves, and had initiated the flowering, with buds already visible.

6.2.10 *KfCDF2*'s impact on reproductive success

To determine the impact of perturbed *KfCDF2* levels on the reproductive success of the transgenic lines, leaves of the same developmental age were removed from each transgenic plant and the wild types. They were left in dry trays in the greenhouse which promotes the development of adventitious leaf margin plantlets, which is the main viable method of reproduction for *K. fedtschenkoi* due to its failure to form viable seeds following fertilisation of flowers. Both the number of notches on each leaf and the number of plantlets produced by each leaf was counted, and these values allowed the calculation of a ratio for the number of notches available per plantlet produced.

6.2.10.1 The impact of *KfCDF2* overexpression on leaf plantlet production

KfCDF2_FL_13A showed no difference in plantlet production compared to the wild type, whereas *KfCDF2_FL_14C* produced more plantlets per leaf notch (Fig. 6. 39; P = 0.0008). As more plantlets were produced per leaf margin notch in line *14C*, it is possible that this line would achieve a greater reproductive success than the wild type, although to test this fully, the plantlets would have to be grown on and it would have to be demonstrated that they could all establish sufficiently well in the next generation to produce leaves large enough to themselves develop leaf margin plantlets.

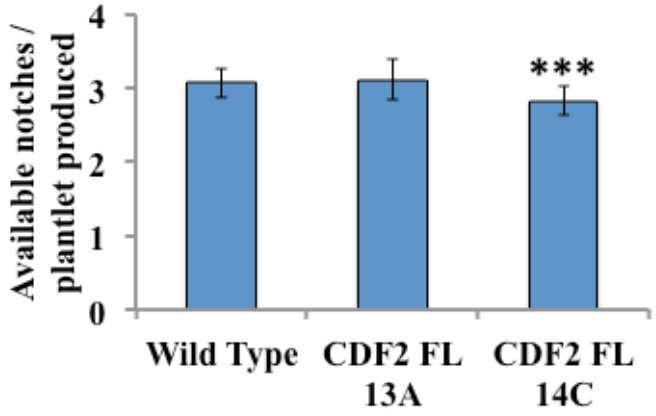

Figure 6.39. Impact of over-expression of *KfCDF2* on reproductive output in the form of leaf margin plantlet formation (For WT n = 39 leaves, 13A n = 18 leaves, for 14C n = 22 leaves). Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = P ≤ 0.05, ** = P ≤ 0.01, *** = P ≤ 0.001, **** = ≤ 0.0001. Stars above each column represent significant differences to the corresponding WT value.

6.2.10.1 The impact of *KfCDF2* RNAi knockdown on leaf plantlet production

For the *KfCDF2_RNAi* lines, both showed a significant decrease in the number of notches required to produce a plantlet (Fig. 6.40). Thus, both RNAi lines were able to produce more plantlets per leaf margin notch compared to wild type ($P < 0.0001$).

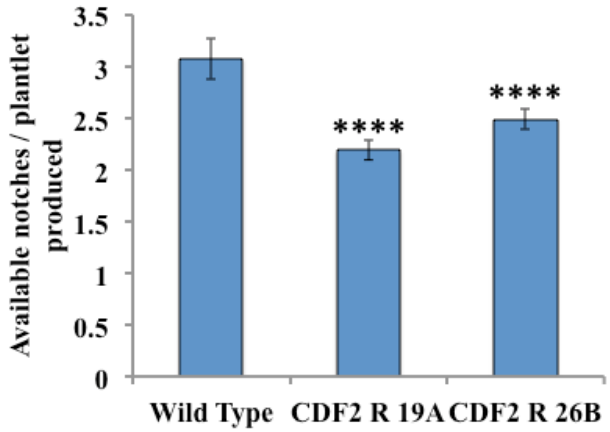

Figure 6.40. Impact of RNAi knockdown of *KfCDF2* on reproductive output in the form of leaf margin plantlet formation (For WT n = 39 leaves, 19A n = 19 leaves, for 26B n = 38 leaves). Tukey HSD was used to calculate significant differences, and stars (*) are used to represent the results; * = $P \leq 0.05$, ** = $P \leq 0.01$, *** = $P \leq 0.001$, **** = ≤ 0.0001. Stars above each column represent significant differences to the corresponding WT value.

6.3 Discussion

This chapter has focused on the impact of transgenic perturbation of the transcript levels of *KfCDF2* on phenotypes associated with CAM, including its control by the circadian clock and the operation of the central circadian clock itself. As described in the preceding chapters for *KfMYB439* and *KfCIB1*, two lines that were successfully confirmed to over-express *KfCDF2,* and two RNAi lines that were confirmed to have a down-regulation of the level of *KfCDF2* transcripts were chosen for detailed characterisation. Again, varying levels of over-expression and knock-down were chosen in order to allow any dosage-dependent effects of the level of expression of *KfCDF2* to be studied. The four lines: *KfCDF2_FL_13A, KfCDF2_FL_14C, KfCDF2_RNAi_19A* and *KfCDF2_RNAi_26B* were subjected to various molecular, biochemical and physiological measurement in order to determine the role, if any, of *KfCDF2* within the circadian optimisation of CAM.

6.3.1 Bioinformatic characterisation of *KfCDF2*

Despite *AtCDF2* (At5g39660) being the closest orthologous sequence to *KfCDF2* in the Arabidopsis genome, when the two encoded amino acid sequences were aligned, they were found to only share 38.8% pairwise identity across the whole protein. The two genes shared a high level of identical amino acids in their DOF and activation domains (around 90% similarity; Fig. 6.2). This suggests that *KfCDF2* has diverged substantially since its last common ancestor with Arabidopsis, especially considering that a highly conserved gene, such as the gene encoding the CAM PEPC in *K. fedtschenkoi,* share 80 % identical amino acid residues with the closest ortholog in the Arabidopsis genome. This rapid and wide divergence of the two orthologous *CDF2*

genes, *KfCDF2* and *AtCDF2*, supports the proposal that *KfCDF2* may carry out different and/ or supplementary functions and roles compared to its Arabidopsis ortholog. The circadian coordination of CAM was therefore a entirely plausible role for *KfCDF2* based on its strong induction in CAM leaves relative to C_3 leaves (Chapter 3).

The DOF domain was the most highly conserved region shared between the two *CDF2* genes, whereas the activation domain showed a greater level of sequence divergence. The changes in the activation domain may be important for the neofunctionalisation of *KfCDF2* for a role in the daily coordination of CAM in *K. fedtschenkoi*. The activation domain would therefore be a key target for mutagenesis work in future, plus it will be important to identify any interecting partners that bind to the activation domain of *KfCDF2*. Furthermore, CDFs have been shown to act redundantly alongside their closely related paralogous pairs many times in other species and all share a conserved DNA binding site (Yanagisawa *et al.,* 1993; Fornara *et al.,* 2009). To investigate this, *KfCDF2* and a closely-related *CDF* TF found in the *K. fedtschenkoi* genome, namely *KfCDF3,* were aligned and compared. These two sequences showed 76.4 % pairwise identity over the whole protein, and so could well be regulated and act in similar ways. Their expression profiles were also found to be very similar using the existing SOLiD RNA-seq dataset for C_3 and CAM leaves of *K. fedtschenkoi* (Appendix: S6.3). This could therefore affect results for the RNAi lines that were generated here, as the hairpin RNA binary construct introduced into the transgenic RNAi lines may not have been able to silence the transcripts of the, potentially redundant, *KfCDF3* gene. There will also be many more DOF TF genes in the *K. fedtschenkoi* genome, as over 30 DOFs have been found in both Arabidopsis and Rice (Yanagisawa 2004).

Previously, other transgenic lines have been created in *K. fedtschenkoi* targetting known CAM-associated or central clock-associated genes, including *NAD-ME RNAi,* and *PPDK RNAi* (Dever *et al.,* 2015), plus a *PPCK RNAi* (Boxall *et al.,* 2015/16 in preparation), and a *TOC1* over-expresser (Dall'omo, 2011). An Illumina RNA-seq quantitative whole transcriptome datasets was previously generated for CAM leaves from each these transgenic lines using samples collected at 4 h intervals over a 12:12 light: dark cycle (Boxall, Dever and Hartwell, unpublished data). This RNA-seq dataset enabled the mis-regulation of *KfCDF2* to be examined in each of these transgenic lines, which have known perturbations in their daily cycle of CAM (Fig. 6.3).

In the *TOC1* over-expresser line, *KfCDF2* transcript abundance was reduced by around 50 % in the dark. TOC1 has been shown to function as a general transcriptional repressor, which could suggest this gene may be involved in a circadian output process regulated by TOC1 (Gendron *et al.,* 2012). Another possibility is that the changes in the core clock caused by the constitutive expression of TOC1 in this line lead to pleiotropic effects on a myriad of processes including the regulation of *KfCDF2*. Therefore more work needs to be conducted in this area.

The level of transcripts for *KfCDF2* also fell in the CAM leaves of the *rPPCK1_3* RNAi line that fails to phosphorylate PEPC in the dark period due to failure to switch on the circadian clock controlled PPCK activity in the dark. DOF TFs have already been shown to directly bind to *PEPC* promoters (Yanagisawa *et al.,* 1998), and so may also bind to other genes involved in CAM. Furthermore, the *rPPCK1_3* line showed much less CO_2 fixation due to more rapid inhibition of PEPC by malate, which is also seen in *KfCDF2* overexpressing and knockdown lines. However, the

regulation of *KfPPCK1* transcripts changed little in the *KfCDF2* transgenic lines (Figs. 6.4 to 6.7) suggesting that the observed reductions in the phosphorylation state of PEPC detected on the immunoblots may be due to some form of metabolic regulation of PPCK activity, such as the known inhibition of PPCK by malate through a substrate interaction with its target protein PEPC.

There was also a large reduction in *KfCDF2* transcript abundance in the *rNAD_ME1* line (Fig. 6.3), which again suggests CAM in some way feeds back and affects the expression/ RNA stability of the *KfCDF2* gene. Alternatively the changes in metabolism in the *rNAD_ME1* line may feed back to alter temporal regulation within the circadian clock, which can then modulate *KfCDF2* transcript abundance. By contrast, for the *rPPDK* RNAi line which has a greater than 90 % reduction in the activity of PPDK, there was only a small reduction in the transcript level for *KfCDF2* relative to the wild type. The *rPPDK* line had similar changes in CAM and CAM associated metabolites to the *rNAD_ME1* line consistent with the fact that it catalyses the next sequential step after NAD-ME during malate decarboxylation in the light (Dever *et al.,* 2015). However, the *rPPDK* line did manage to maintain circadian rhythms in the transcript abundance of core clock genes in LL conditions, whereas the *rNAD_ME1* line displayed arrhythmia for *KfTOC1* transcript rhythms in LL (Dever *et al.,* 2015), suggesting that the core CCA1/LHY-TOC1 oscillator rapidly collapsed to arrhythmia in the *rNAD_ME1* line but not in the *rPPDK* line. Circadian oscillations of *KfTOC1* and *KfCCA1* transcript levels also collapse rapidly to arrhythmia under LL conditions in *rPPCK1_3* and the *TOC1-OX* line, which also displayed large reductions in the transcript abundance of *KfCDF2* (Fig. 6.3). Thus, all of the transgenic lines that display arrhythmia for core clock gene transcripts in LL conditions also displayed a marked reduction in the transcript abundance of *KfCDF2*

(Fig. 6.3), suggesting that the core clock perturbations that lead to LL arrhythmia in these lines also has a knock-on effect on the light/ dark regulation of *KfCDF2* transcript levels.

6.3.2 *KfCDF2*s impact on gene expression

Both over-expresser lines and both RNAi knockdown lines identified in Chapter 3.2.4, have been confirmed as having functional transgenes, with *KfCDF2_FL_13A* being the strongest over-expresser line and *KfCDF2_RNAi_26B* showing the largest reduction in transcript. As previously mentioned, DOF transcription factors show many redundancies, and so therefore double and triple mutants may be a next step, to investigate the CDF family as a whole, and its impact on CAM. It has also been demonstrated Arabidopsis that if one CDF was overexpressed, for example *AtCDF3*, it can lead to an increase in the expression of other DOFs (Yanagisawa *et al.,* 2004). The diversity of the DOF family in *K. fedtschenkoi,* and the regulation of each gene in the family, will be a fruitful area for future study as it will aid the identification of potential partners for *KfCDF2,* which may work with this gene in order to achieve the correct circadian control of CAM, and other clock controlled processes in the leaves including the induction of florigen in response to short days (see Chapter 7). There has also been recent research which defined role of CDF2 in the 'GI-CDF module' for regulating gene expression in Arabidopsis (Fornara *et al.,* 2015).

As well as the changes to *KfCDF2* expression level in the transgenic lines, a number of other CAM and clock-associated genes were also found to be mis-regulated relative to the wild type, although only small perturbations in the transcript abundance of other genes were detected. The transcript oscillations of many CAM-associated genes showed minor perturbations in both the over-expressing and RNAi knockdown

lines, suggesting that *KfCDF2* may play a role in the output from the clock as originally hypothesised, due to the transgene effects being detected for so many CAM genes. Furthermore, changes in *KfCCA1* and *KfTOC1* were also detected, especially for *KfTOC1* in *KfCDF2_FL_14C* and *RNAi_19A* (Figs. 6.5 & 6.6). This could indicate that, in addition to *KfCDF2* functioning in the clock output pathway, it may also feed back to and interact with the clock. This notion is also supported by the *TOC1_OX* sequencing data, which showed that *KfCDF2* transcript levels were repressed strongly when *TOC1* was over-expressed strongly from the 35S promoter. It is possible that *KfCDF2* does feedback to the clock, as numerous DOF TFs have been shown to play roles in light signalling. For example, in Arabidopsis DOFs have been shown to function in phytochrome A signalling, which in turn feeds light input signals into the central oscillator (Tepperman *et al.,* 2001).

Subtle changes in the transcript abundance profiles of the CAM genes *KfPEPC* and *KfPPDK* were detected in the *KfCDF2_FL_14C* line and also both *KfCDF2_RNAi* lines (Figs. 6.5, 6.6 & 6.7). As most DOF transcription factors bind to very similar sequences in their target promoters, and maize *Dof1* was found to bind to the promoter of the C_4 PEPC gene, then it is possible that *KfCDF2* may be able to bind directly to the promoters of these two key CAM genes, as similar binding activity has already been confirmed for the equivalent promoters in maize (Yanagisawa 2000). It is also possible that *KfCDF2* binds to the promoters of other CAM genes as well, therefore more bioinformatics work is required in this area to determine binding motifs in CAM-associated gene promoters.

DOF transcription factors have also been shown to control carbon metabolism, which is interesting, as starch metabolism and sugar transporter genes also showed changes in their transcript abundance oscillations under LD cycles in some of the *KfCDF2*

lines (Figs. 6.4 – 6.7; *GPT2, MEX1, GWD*). It should be noted though that, as mentioned before, a change in one *KfCDF*, especially over-expression, could potentially lead to changes in other *KfCDF*s, which may in turn lead to widespread knock-on effects spanning a number of biological processes.

It should be noted that *AtCDF2* has been shown to be regulated by a number of different clock proteins in a post-transcriptional manner, such as the clock components GI and ZTL. Thus, in order to link changes in other genes to *KfCDF2* transcript levels, protein level quantification should be undertaken to confirm that gene expression changes reported here lead to protein abundance changes (Imaizumi *et al.,* 2005; Fornara *et al.,* 2009; Song *et al.,* 2014).

KfCDF2_FL_14C and *KfCDF2_RNAi_26B* were chosen to be investigated under free-running, constant LL conditions in order to investigate in more detail any perturbations in circadian oscillations. *FL_14C* was chosen over *FL_13A* because line *13A* proved difficult to establish successfully, and *14C* still showed a large variety of effects on gene expression rhythms, despite not over-expressing *KfCDF2* as strongly as line *13A*.

In LL, *FL_14C KfCDF2* showed a phase advance, peaking four hours earlier for the first two full oscillations (Fig. 6.8). By the third oscillation though, there was a large reduction in amplitude, such that the wild type and line *KfCDF2_FL_14C* had very similar patterns of transcript abundance for *KfCDF2* (Fig. 6.8). The over-expression of *KfCDF2* led to small reductions in the transcript levels for *KfCCA1, KfTOC1* and *KfPPCK1*, suggesting that *KfCDF2* in some way contributed to the circadian rhythms in the transcript levels of these genes.

With *KfCDF2_RNAi_26B* in LL, *KfCDF2* transcript oscillations were arrhythmic from 30:00 LL to 46:00 LL, but afterwards returned to a rhythmic pattern, but at

lower transcript level compared to the wild type (Fig. 6.9). This arrhythmia of *KfCDF2* for over 24 h also impacted on the oscillations in *KfCCA1* and *KfTOC1*, which did not show changes in transcript abundance during this period (Fig. 6.9; first black and white dotted boxes). Once *KfCCA1* and *KfTOC1* transcript oscillations became rhythmic again, both showed a 4 h phase delay in the timing of their peak. This suggests changes in circadian clock signalling have occurred, as the transgenic line was anticipating the subjective time differently to the wild type. This again suggests *KfCDF2* may feed back signals to the clock, similar to the way in which other DOFs have been found to affect the phytochrome A signalling pathway which inputs light signals into the clock (Somers *et al.,* 1998; Tepperman *et al.,* 2001; Yang *et al.,* 2010). Furthermore, CDFs in Arabidopsis have been shown to function downstream of GI and also interact antagonistically with GI in a variety of processes, and so this could be the way that *KfCDF2* interacts with the core clock (Fornara *et al.,* 2015).

6.3.3 Gas exchange rhythms

Both *FL* over-expresser lines displayed delays in CO_2 fixation under LD conditions (Figs. 6.10 and 6.11). When *FL_13A* opened its stomata at the beginning of the dark period, a large amount of CO_2 was released (Fig. 6.10A & B), suggesting there may still have been a high internal partial pressure of CO_2 present within the leaf. This is consistent with the fact that phase IV was virtually absent in this line (Fig. 6.10). Despite the start of dark period CO_2 fixation being delayed relative to the wild type in *FL_13A*, this line was still able to fix large amounts of CO_2 during the dark period, and overall managed to fix more CO_2 than wild type (Table. 6.2). *FL_14C* on the other hand did not fix as much CO_2 as the wild type during the dark period, and as a

result carried on fixing CO_2 into the light period (Fig. 6.11B). *KfCDF2_FL_14C* fixed 25 % less CO_2 in the dark period compared to the wild type, even though its period of dark CO_2 fixation was extended into the light period (Table. 6.2).

In LL conditions, both lines became arrhythmic after a few days. It took 60 h for *FL_13A* to become arrhythmic and 96 h for line *FL_14C* (Figs. 6.10 and 6.11). This therefore suggested that higher expression of *KfCDF2* correlated with a more rapid collapse of the CO_2 exchange rhythm to arrhythmia. The fact that the CO_2 exchange rhythms of these *FL* lines became arrhythmic under LL conditions supports the proposal that *KfCDF2* feeds back to the clock, or is coupled to the clock in some way, as the *McTOC1_OX* over-expressing lines of *K. fedtschenkoi* also displayed a very similar collapse of the CO_2 exchange rhythm to arrhythmia (Dall'omo 2011).

The two RNAi lines showed varying results. *RNAi_26B* showed similar results in LD to *KfCDF2_FL_14C*, as there was a delay in CO_2 fixation at the beginning of the dark period, and a lower amount of CO_2 was fixed during the dark period (Fig. 6.13). This line also went arrhythmic after 120 h in LL, suggesting that repression of *KfCDF2* did not have as big an effect on the clock as over-expression did. *KfCDF2_RNAi_19A* on the other hand, showed phase IV gas exchange before the wild type (Fig. 6.14). This line also fixed more CO_2 during the dark period, and also did not become arrhythmic in LL. This may be because *RNAi_19A* was the line with the lower level of reduction in the level of *KfCDF2* transcripts; only 1.3-fold in the light and 1.2-fold during the dark, and so this could suggest that a small reduction in *KfCDF2* levels can be tolerated, whereas the larger changes in line *KfCDF2_RNAi_26B* (2.2-fold in the light and 3.2-fold in the dark) were sufficient to cause larger impacts on CO_2 fixation in LD cycles and the rhythm of CO_2 fixation in LL.

6.3.4 Metabolite levels

Metabolite levels of malate, starch and sugars were quantified, as these metabolites are needed for, and produced during, the efficient functioning of CAM. Both over-expresser lines showed a clear trend of decreased malate accumulation at the end of the dark period and increased levels at the end of the light, which therefore resulted in reduced malate turnover over the 24 h period (Fig. 6.14). *KfCDF2_RNAi_19A* also showed this same result, whilst *RNAi_26B* showed no difference to wild type (Fig. 6.18). These results are not consistent with the CO_2 fixation data, as both *FL_13A* and *RNAi_19A* fixed more CO_2 than the wild type in the dark, which should lead to more stored at dawn. Furthermore, *KfCDF2_RNAi_26B* fixed much less CO_2 than the wild type, so it is difficult to understand how it was able to accumulate a similar amount of malate relative to the wild type. However, the anomaly between these results may be a sign that the *KfCDF2* lines undertake more futile cycling in the dark; for example performing simultaneous CO_2 fixation and malate decarboxylation. However, for the CO_2 fixation measured by the IRGA system to be higher than the wild type, the extra CO_2 that was respired out of malate due to the futile cycle would have to fail to make it back out of the leaf through the stomata. Alternatively, it could be that the extra dark CO_2 fixation in these lines did not accumulate as malate, but instead the extra fixed carbon was redirected to another storage metabolite. Comprehensive metabolite profiling of these transgenic lines over the light/ dark cycle using CAM leaves could resolve this possibility in the future. In addition, feeding $^{13}CO_2$ or $^{14}CO_2$ to the lines in classical pulse-chase experiments would also allow the further elucidation of what these lines do with the CO_2 that they fix during the dark period.

At the end of the dark period, the *FL* lines had turned over more starch than the wild type and this correlated with higher levels of soluble sugars in both lines (Figs. 6.15 &

6.17). For the RNAi lines, there was a small decrease in soluble sugars in the light, but there was an increase in dark sugar levels (Fig. 6.21). DOF factors have previously been shown to play roles in carbohydrate metabolism, with effects on both starch and sugar levels (Yanagisawa *et al.,* 1998) . When the *Dof1* gene from maize was expressed in potato, it increased starch accumulation dramatically (Kisaka *et al.,* 2007), and similar results were also found for another *DOF* gene, *SRF1* in sweet potato (Tanaka *et al.,* 2009). In Arabidopsis over-expression of *Dof1* has also been shown to lead to decreases in glucose (Yanagisawa *et al.,* 2004). Although this does not happen in the transgenic lines studied here, DOFs have been shown to play antagonistic roles with one another, and so this could well be the case for *KfCDF2* in *K. fedtschenkoi*, where *KfCDF2* over-expression led to an increase in sugars rather than a decrease.

The increase in soluble sugars in several of the *KfCDF2* transgenic lines could be one reason why many of these lines became arrhythmic in LL, due to sugar levels feeding back to the clock (Haydon *et al.,* 2013).

6.3.5 Protein levels and enzyme activity

For both *FL* over-expresser lines and the *RNAi_26B* line, immunoblots were performed which allowed the steady state abundance of PEPC protein and phosphorylated-PEPC to be established (Figs. 6.22 and 6.23). PEPC levels were approximately stable throughout the 12:12 light/ dark time course, suggesting that the small changes in PEPC transcript levels in all three lines did not result in detectable changes in the amount of PEPC protein in the leaves (Figs. 6.22A & 6.23A). When looking at the pattern of phosphorylated-PEPC, all three lines showed a decrease in phosphorylation at the start of the dark period (Figs. 6.22B & 6.23B). Thus, as CO_2

was fixed by all three lines in the early part of the dark period, it was likely that dephosphorylated-PEPC was starting to be inhibited by malate accumulation. This would therefore explain the IRGA data for these three lines, which showed a delay in the rise in CO_2 fixation at the start of the dark period.

To study these possibilities in more detail, both PEPC assays and PEPC malate sensitivity assays were conducted (Figs. 6.24 & 6.25 and Tables 6.4 & 6.5). The PEPC assays suggested that all four transgenic lines showed in increase in PEPC activity when the values were expressed per mg total protein (Fig. 6.24). However, when the PEPC activity values were expressed per my of chlorophyll, only *KfCDF2_RNAi_19A* showed an increase in the total extractable activity of PEPC (Figs. 6.24 & 6.25). This is interesting, as the over-expression of maize *Dof1* in Arabidopsis has been shown to lead to an increase in amino acid content in plants (Yanagisawa 2000), whereas in *K. fedtschenkoi* with perturbed expression of *KfCDF2*, there may be a general decrease in total protein levels relative to the level of PEPC, leading to an apparently higher activity of PEPC compared to wild type when the values were expressed per mg of protein. Further work to examine the total amino acid and protein levels in the CAM leaves of these *K. fedtschenkoi KfCDF2* lines should provide valuable further insights into these unexpected results.

The increase in PEPC activity in line *RNAi_19A* even when expressed per mg of chlorophyll was supported by the CO_2 exchange data, and also by the increased malate accumulation at the end of the dark period in this line (Figs. 6.12 and 6.18).

With the PEPC malate sensitivity assays, all four lines showed an increased sensitivity to malate inhibition in both the light and dark (Tables 6.4 & 6.5). This goes some way to explaining why decreases in malate levels were detected for *KfCDF2FL_13A, FL_14C* and *RNAi_26B*, as if the PEPC in these plants was more

sensitive to malate inhibition than the wild type during the dark period, then the enzyme would be inhibited at a lower level of nocturnal malate accumulation, and thus the total malate accumulation by dawn would be diminished relative to the wild type.

6.3.6 Performance of the *KfCDF2* transgenic lines under drought-stress

As proper functioning of CAM enables high WUE in water limited conditions, it was important to investigate the performance of the KfCDF2 lines under drought-stress conditions. Six-month-old plants were either left in well-watered conditions (watering every day) or were subjected to drought-stress by with-holding watering for 31 days. At the end of the drought treatment, the above- and below-ground material was harvested, plus LP6 were collected for pigment and malate quantification.

The reduction of *KfCDF2* in the RNAi lines did not have a large effect on the response of the plants to drought. This may be due to redundancy of *KfCDF2* with other *KfCDFs* (Fig. 6.30). From experiments in Arabidopsis, a quadruple *cdf1 cdf2 cdf3 cdf5* mutant line showed decreases in genes essential for drought-stress tolerance, such as *EARLY RESPONSIVE TO DEHYDRATION7* and *10* (*ERD7 & ERD10*), leading to a poor performance in drought (Fornara *et al.,* 2015). This again was a function that involved GI, so it would be interesting to determine the levels of GI and the *KfCDF1, 3* and *5* orthologs in the drought-stressed lines studied here.

Looking at the *KfCDF2_FL* over-expresser lines, both lines showed significant fresh and/or dry weight differences to wild type. *KfCDF2_FL_13A* showed a reduction in fresh weight and dry weight in well-watered condition, whilst *FL_14C* showed an increase in the fresh and dry weight for above-ground tissues under drought-stress (Fig. 6.26). In terms of the root biomass below ground, *KfCDF2_FL_13A* again

showed a decrease in fresh and dry weight. There was also an increase in root fresh weight for *KfCDF2_FL_14C*. For the root dry weights, for both well-watered and drought-stressed conditions, there was a decrease compared to the wild type. Both lines also showed a significant increase compared to wild type in percentage moisture for both shoots and roots (Fig. 6.26E & F). Corrales *et al.* (2014) found similar results when overexpressing *CDF1* and *CDF3* in tomatoes, which showed an increase in drought tolerance by increasing survival in long-term drought, and also led to an increase in root length.

When total leaf anthocyanin levels were quantified, all lines showed a large increase in anthocyanins between well-watered and drought-stressed conditions. The *FL* over-expresser lines showed no difference in anthocyanin levels compared to the wild type (Fig. 6.27), whilst the RNAi lines produced significantly less anthocyanin pigments than the wild type (Fig. 6.31). This could be because other essential genes involved in the drought-stress response may have been down regulated in these lines. For example, the *ERD* genes shown to be down regulated in Arabidopsis (Fornara *et al.,* 2015). The regulation of these genes and other known drought-responsive genes in these *K. fedtschenkoi KfCDF2* transgenic lines would merit further investigation to determine whether the *CDF* genes in *K. fedtschenkoi* also play a role in the drought-stress responsive signalling pathway in this species.

Several photosynthetic pigments were also quantified, namely ChlA and B and carotenoids. Under well-watered conditions, the *FL* over-expresser lines both had a significant reduction in ChlA, whereas the RNAi lines showed no difference (Figs. 6.28 & 6.32). *KfCDF2_FL_13A* also showed a reduction in ChlB in drought-stressed conditions. *KfCDF2_FL_13A* also had reduced levels of carotenoids under drought-stress conditions, and again, the RNAi lines showed no differences.

Malate levels were quantified to determine whether the plants were still performing full CAM after the 31 day of drought-stress treatment, or whether they had started to perform CAM-idling, wherein they keep their stomata closed throughout the light/ dark cycle and recapture respired CO_2 via PEPC in the dark period behind closed stomata. In all four *KfCDF2* transgenic lines, malate accumulation in the dark period was significantly lower than the level detected for the wild type, and there was also a much smaller decrease in daily turnover of malate in the transgenic lines compared to the wild type (Figs. 6.29D & 6.33D). This suggested that CAM-idling may have been occurring in these lines. This could be confirmed by using an infra-red gas analyser such as a LICOR 6400XT system in order to measure stomatal conductance. In Arabidopsis, genes known to be regulated by *CDF2*, including *SOC1, CO* and *FT*, which are all involved in the flowering pathway, have also been shown to impact on stomatal control and development if repressed, leading to the opening of stomata at incorrect times (Kimura *et al.,* 2015). It would therefore also be of interest to investigate the regulation of stomatal conductance in these lines over the LD cycle in well-watered conditions, in addition to further detailed analysis of the stomatal physiology of these lines under drought-stress

The LICOR-6400XT gas exchange analysis system could also be used to determine how many days under drought it takes for each line, including the wild type, to reach CAM-idling, as this system would allow for the accurate determination of point at which the plants switched to keeping their stomata closed throughout the 24 h cycle.

6.3.7 Performance in long-term drought

From the 31-day drought experiment, it became apparent that some of the *KfCDF2* lines, especially *KfCDF2_FL_13A,* had an increase in above-ground moisture content

compared to the wild type in well-watered and drought-stressed conditions (Fig. 6.26). This could be a direct result of the decrease in phase IV of CAM in this line, as seen in the gas exchange data (Fig. 6.10), which could result in less water being lost. Alternatively, the perturbation of the level of *KfCDF2* expression in these lines may in some way lead to an increase in the water storage capacity of the cells, possibly via an increase in cell size or vacuolar water storage capacity. With these considerations in mind, it was decided to study the succulence of each *KfCDF2* transgenic line and the wild type under both well-watered conditions and prolonged drought-stress conditions. Both the wild type and the *KfCDF2* lines all continued to increase in total above ground succulence when maintained under well-watered conditions (Fig. 6.34). The transgenic lines showed a higher rate of increase in leaf succulence than the wild type over the 90 days.

Under drought conditions, all of the lines were able to maintain a higher level of succulence than the wild type throughout the 90 days of water with-holding (Fig. 6.35). Under drought, the two-over-expresser lines were retained the highest level of leaf water per unit area at the end of the 90 days, with *FL_13A* showing the highest value, which correlated with the higher level of *KfCDF2* over-expression in line *FL_13A*. This phenotype may also have been present in the closely related species *K. laxiflora*, as the recently generated *KlHA-KfCDF2* over-expresser lines, also have small but thick leaves, based on visual observations, and have already been found to perform better under drought-stress than the wild type *K. laxiflora* (Fig. 6.38). Related findings have been reported previously when CDFs were over-expressed in tomato plants, where over-expression of *CDF3* led to increased sucrose levels in the whole plant and also an increase in amino acids, amongst other metabolites, which are often accumulated in the plant during times of stress (Corrales *et al.,* 2014). The *KfCDF2*

over-expresser lines shared some similar phenotypes, for example increased dark period sugar levels, plus, possibly, changes to amino acid levels in leaves (based on the odd results when PEPC activity was expressed relative to total leaf protein levels), then these plants may already be primed for drought. This could also explain why they have increased in succulence, as by the increase in soluble sugars could act as compatible solutes, which act as an osmoticum that helps to maintain a high water content in the leaf cell vacuoles.

6.3.8 Reproductive output of leaf margin plantlets

Using detached leaves removed from each transgenic line and the wild type, the ratio of plantlets per leaf notch was calculated to quantify the ability of each lines to reproduce via adventitious leaf margin plantlets. Leaf plantlets can be considered the main reproductive output in a species such as *K. fedtschenkoi* which has lost the ability to set viable seed (Garces *et al.,* 2007). *KfCDF2_FL_13A* had a similar ratio plantlets to leaf notches to the wild type, but showed a bigger variation plantlets per leaf notch, whilst the other three lines all showed a slight decrease in notch to plantlet ratio. This indicated that these lines produced more plantlets and so had the potential to increase their reproductive output relative to the wild type (Figs. 6.39 & 6.40). DOF transcription factors have been found to regulate plantlet growth previously, especially in terms of influencing seed germination in various species including Arabidopsis, maize and potato, suggesting that changes in the levels of certain DOFs may lead to more efficient plantlet production (Gabriele *et al.,* 2010; Venkatesh *et al.,* 2015).

6.4 Conclusions

The experimental evidence presented in this chapter support the proposed role of *KfCDF2* in the circadian optimisation of CAM, and also suggests that *KfCDF2* may be able to feed back to influence the central clock.

KfCDF2 mis-expression had a widespread, but relatively small, effects on the transcript abundance of a range of screened CAM and clock genes measured over a 12:12 light/ dark time course. These data suggest that *KfCDF2* probably does not regulate any of the measured CAM genes directly. The LL transcript abundance data showed changes in the timing of key clock- and CAM-associated genes, and *KfCDF2* transcript levels have also been found to be reduced in transgenic lines over-expressing the core clock gene *TOC1*, suggesting a link between *KfCDF2* and the clock.

CO_2 fixation in all of the *KfCDF2* transgenic lines also showed phase shifts compared to the wild type; specifically in terms of the timing of the start of the subjective-dark burst of CO_2 fixation. Furthermore, the larger perturbations in the expression level of *KfCDF2* did affect the circadian regulation of CAM in LL, as after a few days under LL conditions, the CO_2 fixation rhythm became arrhythmic except 19A, whereas the wild type continued to oscillate robustly. Similar CO_2 fixation arrhythmia had already been recorded in transgenic *TOC1-OX* lines of *K. fedtschenkoi*, and also in several other transgenic lines in which key CAM genes, including *KfPPCK1* or *KfNAD-ME_b1,* were silenced using RNAi binary constructs (Dall'omo 2011; Dever *et al.,* 2015).

Changes in *KfCDF2* expression also had an impact on the levels of key pools of metabolites associated with CAM. Both *KfCDF2_FL* lines plus line *RNAi_19A* showed decreased malate levels at the end of the dark, and an increased level at the

end of the light relative to the wild type. This reduction in the nocturnal accumulation and diurnal turnover of malate correlated well with the measured decrease in phosphorylated-PEPC, and the reduction in the apparent K_i of PEPC for malate compared to wild type. From sugar levels, all *KfCDF2* transgenic lines showed large increases in total soluble sugar levels at the end of the dark period.

Under drought-stressed conditions, all four transgenic lines showed higher values for their percentage moisture compared to the wild type, with the *KfCDF2_FL* lines showing the largest values. To further investigate this, all of the *KfCDF2* transgenic lines and the wild type were subjected to either a 90 day drought-stress treatment by with-holding water, or maintained well-watered throughout. In well-watered conditions, all of the *KfCDF2* transgenic lines increased their leaf succulence at a faster rate and were more succulent than the wild type. Under drought-stress, after 90 days leaves of all of the transgenic lines were more succulent than the equivalent leaves from the wild type, with *FL_13A*, the strongest over-expresser line, showing the highest succulence value. Furthermore, after re-watering the remaining plants, both FL lines fully recovered within a week and shed no leaves, whereas both RNAi lines and wild type lost leaves during their recovery from the 90 day drought-stress treatment.

Finally, when reproductive productivity of the *KfCDF2* transgenic lines was measured, there was generally an increase in plantlet productivity, with fewer notches being required to produce a plantlet. The success of these plantlets in establishing in soil, and growing on sufficiently for them to produce mature leaves that are themselves competent to produce leaf margin plantlets, still needs to be tested.

This work aimed to define the role, if any, of *KfCDF2* in the temporal coordination and optimisation of CAM in response to the central circadian clock. It has often

proved challenging in other species to define the role of a single member of a diverse and multi-membered transcription factor family. Often the whole family, or at least a key sub-family needs to be investigated due to problems such as gene redundancy (Fornara *et al.,* 2009), antagonism within the family (Papi *et al.,* 2000), and also interactions between them. Furthermore DOF TFs have such a wide variety of functions that they are involved in most processes at some point, either as a repressor and/or an activator (Papi *et al.,* 2000).

Furthermore, DOF TFs also interact with other proteins, from a variety of other families, including other TFs and chaperone proteins (Albani *et al.,* 1997; Bustin 1999; Diaz *et al.,* 2002). These interactions function to moderate activity and binding, and DOFs can also moderate the DNA-binding characteristics of other TF families as well. This therefore poses problems when assessing the functions and impact of one TF, as at certain times of the day the other proteins will not be present.

A key area for future research will be to identify *KfCDF2* target sites that are bound by this TF allowing it to directly control the expression of downstream genes throughout the *K. fedtschenkoi* or *K. laxiflora* genome. Chromatin-Immunoprecipitation followed by sequencing (ChIP-Seq) was planned for *KfCDF2* in order to determine exactly where it binds to the genome *in vivo*. However, time constraints prevented this work from being completed. HA-tagged *KfCDF2_FL* over-expression lines were generated in *K. laxiflora*, screened for transgene expression level, and grown up ready for further experiments. Light/ dark time course samples were also collected. Future ChIP-Seq work on these time course leaf samples should help to place *KfCDF2* more firmly into the signalling pathway that mediates the temporal coordination and optimisation of CAM in response to the endogenous

circadian clock by identifying exactly which CAM-associated genes are bound by *KfCDF2* within their upstream regulatory promoter sequences.

Chapter 7 – An investigation of short-day dependent induction of flowering in *K. fedtschenkoi* and its perturbation by transgenic manipulation of the *CYCLING DOF-FACTOR2* transcription factor

7.1 Introduction

The *KfCDF2* over-expressing lines described in Chapters 3 and 6, were found to transition to flowering under long-day 16: 8 light: dark greenhouse conditions, whereas the wild type has never been observed to flower under long-days. It was reported in Chapter 3, that the more strongly over-expressing *KfCDF2* transgenic lines produced flowers in long days, whilst several lines that had lower levels of over-expression only produced hooked meristems; an early sign of the transition of the shoot apical meristem towards becoming a floral meristem. This therefore suggested that *KfCDF2* plays a role in the regulation of flowering time, which is a process that is known to be tightly regulated by the circadian clock. As previously mentioned, the most thoroughly characterised ortholog of *KfCDF2* is called *CDF2* in Arabidopsis, for which many papers have been published. *DOF* TFs, namely *CDF1*, *2*, *3* and *5*, have all been shown to function somewhat redundantly in flowering induction in Arabidopsis (Fornara *et al.,* 2009). The phenotype after genetic manipulation in *K. fedtschenkoi* and Arabidopsis was found to differ, likely because Arabidopsis is a long day flowering plant (LDP), whilst *K. fedtschenkoi* is a short day plant (SDP).

The Arabidopsis *cdf2* mutant lines displayed early flowering compared to wild type in both long and short days (Fornara *et al.,* 2009), whilst constitutive overexpression using the CaMV35S promoter led to a delay in flowering, but only in long days (Yanagisawa 2002). From the initial screening results presented in Chapter 3, *KfCDF2* overexpression was only found to cause early flowering in long days, when these plants normally do not flower, so it was of interest to identify phenotypes when the plants were grown under different day lengths, to investigate whether the *K.*

fedtschenkoi CDF2 gene performs the reverse function relative to the orthologous gene in Arabidopsis.

7.1.1. Flowering time regulation in higher plants

7.1.1.1 Regulation of flowering time in Arabidopsis via the day-length dependent pathway (long day flowering)

Flowering in Arabidopsis can be initiated by four genetic pathways: gibberellin (GA), autonomous, vernalisation and light-dependent (Komeda 2004). The strict regulation of the timing of flowering is believed to be under this range of regulatory pathways because reproduction requires a lot of plant resources, and so plants need to ensure that they flower when the conditions are optimal, both in terms of their own development and the external environment and so flowering has been demonstrated to be a repressible process (Komeda 2004; Suzuki *et al.,* 2006). It has long been established that day length is perceived by mature leaves (Knott 1934), which results in a mobile flower-inducing signal called 'florigen' being produced (Chailakhyan 1936). The florigen then moves from the leaf to the meristem (Zeevaart 1976). In Arabidopsis, the gene *FLOWERING LOCUS T* (*FT*) has since been identified as possessing the properties of a mobile florigen flower-inducing signal (Huang *et al.,* 2005). For FT to be able to transmit this flowering signal the protein must move from leaf companion cells into the phloem and then to the shoot apex (Corbesier *et al.,* 2007; Mathieu *et al.,* 2007). The ER-localised protein FT-INTERACTION PROTEIN1 (FTIP1) enables export of FT from the companion cells to the phloem (Liu *et al.,* 2012). FT is then able to travel to the meristem, where FD is being expressed and the FT-FD complex is then able to induce other genes essential for

flowering (Abe *et al.,* 2005). Therefore after *FT* expression, the gene changes related to flowering are most likely located in and around the meristem.

Monitoring seasonal changes in both the quality and quantity of light is a key method that plant use to determine the time of year. The changing properties of a plant's light environment are interpreted via the circadian clock, and signals generated as a result of changing light quality and quantity can induce flowering via an output pathway from the central clock. *GIGANTEA* (*GI*) is a key component of the clock which also plays a critical role in linking the core clock to the light-dependent flowering signalling cascade that can lead to flowering (Fig. 7.1). GI induces the expression of *CONSTANS* (*CO*), a Zinc finger domain protein. *CO* is also up-regulated by the autonomous pathway, temperature, GA and vernalisation cues (Pose *et al.,* 2012). The strength of this signal is proportional to the earliness of flowering. *CO* is regulated by other Zinc finger domain proteins: *CYCLING DOF FACTORs 1, 2, 3* and *5* (*CDF1,2,3&5*). These are able to repress *CO* and prevent flowering.

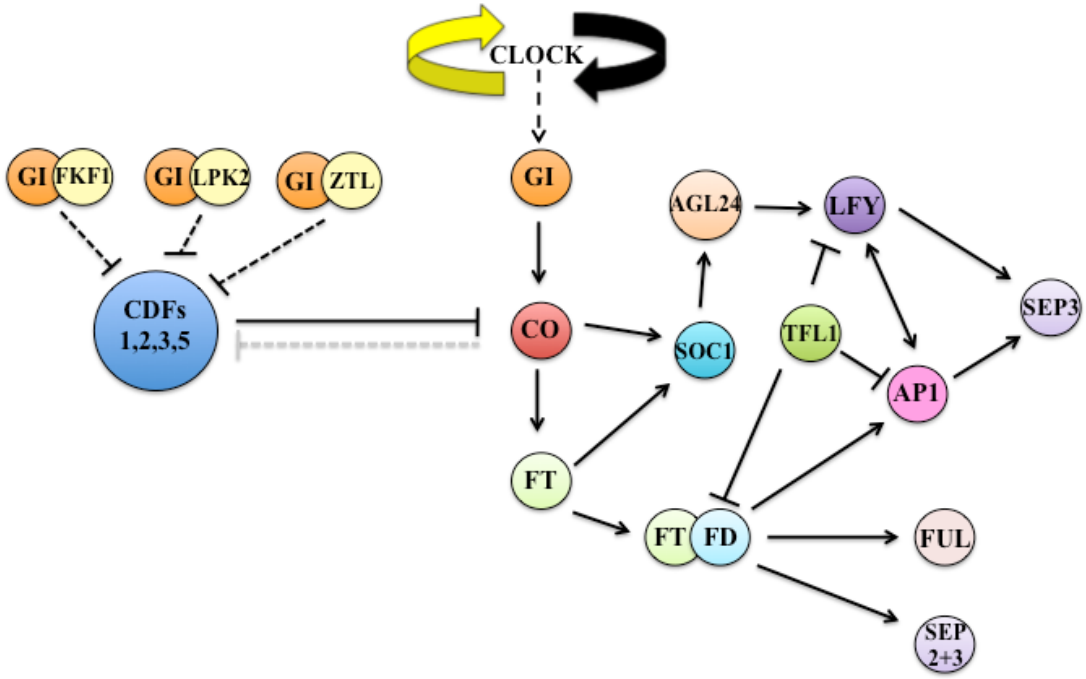

Figure 7.1. Schematic diagram showing an outline of the basic photoperiod-dependent flowering time pathway in the long-day plant *Arabidopsis thaliana*. GI sends the signal from the clock when the light period is long enough to induce flowering. This in turn induces *CO*. CO induces *SOC1* and also *FT*, which then go on to induce many other flowering genes. Once *AP1*, *SEP3* and *FUL* are expressed in the meristem, flowering is initiated. Following the activation of these genes, different classes of genes are activated in different whorls of the developing floral meristem to define and control floral organ development.

In Arabidopsis, once *CO* expression levels reach a threshold that overlaps with the end of the light period in long-day conditions, it is thought to interact with other proteins in order to bind to the promoter of and activate transcription of *FLOWERING LOCUS T* (*FT*) and *SUPRESSOR OF OVEREXPRESSION OF CONSTANS1* (*SOC1*). FT is then able to move through the phloem from the leaf to the meristem, where FT protein interacts with FD. This protein complex can then further activates the expression of the TF SOC1, which then forms a complex with AGAMOUS-LIKE24 (AGL24), which can then promote the expression of *LEAFY* (*LFY*) and *APETALA1* (*AP1*), which are floral meristem identity genes (Yoo *et al.,* 2005). Increased expression of *SOC1* and *AGL24* is often referred to as the key step regulating the

transition from vegetative to inflorescence meristem development (Wellmer *et al.,* 2010).

In later stages of the floral induction process, *SOC1* and *AGL24* are repressed, to prevent reversion of flowering (Lee *et al.,* 2010). As part of the flowering checkpoints, which fine-tune the balance-point between the development of the inflorescence meristem and floral meristem, the FT homolog TERMINAL FLOWER1 (TFL1) acts antagonistically to FT, and so can repress *LFY* and *AP1* expression, and therefore prevent flowering progression. However, once sufficient AP1 has been produced, which produces a strong flowering signal, it can in turn repress *TFL1* and enable flowering to progress efficiently.

Once FUL, SEP3 and AP1 accumulate in the meristem, the inflorescence to floral meristem transition has begun (Teper-Bamnolker *et al.,* 2005).

This high expression of *AP1* also begins the floral meristem development, where sepals, petals, stamen and carpels are formed, by via the ABCDE model of floral organ identity development (O'maoileidigh *et al.,* 2014).

7.1.1.2 Floral meristem identity genes

Proper development of floral organs at the apical meristem is maintained by five classes of floral organ identity genes found in different amounts within the four whorls: A, B, C, D and E (Fig. 7.2). The proteins in class A are AP1 and AP2. They are able to repress C class genes, enabling sepals and petals to form. B class genes are *AP3* and *PISTILLA* (*PI*), which are involved in the formation of petals and stamens, and C class gene *AGAMOUS* (*AG*) is involved in carpel formation. Group D genes: *SEEDSTICK* (*STK*) and *SHATTERPROOF1* and *2* (*SHP1* & *2*) are responsible for the development of ovules within the carpels (Galimba *et al.,* 2015), and group E genes,

SEP1,2,3 and *4,* function as co-regulators in all whorls. The up-regulation of *AP1* signals the start of floral meristem determination. The floral meristem is also maintained by the positive feedback loop between AP1 and LFY (Pose *et al.,* 2012; O'maoileidigh *et al.,* 2014).

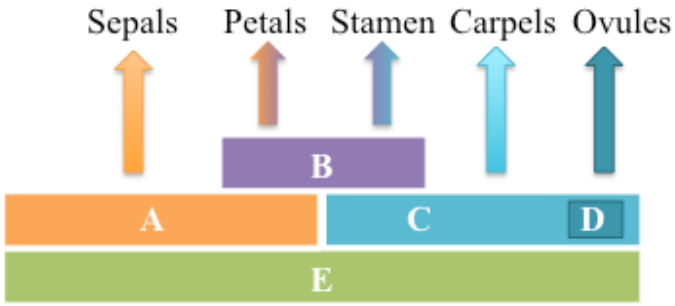

Figure 7.2. Schematic diagram showing the four gene classes involved in floral organ determination. Different interactions between the classes leads to different floral organs being produced by different parts of the floral meristem.

At the beginning of flower development, defined as stages one and two at the meristem, the primordium proliferates without differentiation, which is achieved by the silencing of *SEP3* by SOC1, ALG24 and SHORT VEGETATIVE PHASE (SVP), which interacts with TFL1. At stage three, differentiation begins, with formation of the outer whorl: sepals. At this point *SEP3* expression is promoted by AP1 and LFY, which can then go on and activate B and C class genes in the three inner whorls, which in turn leads to *SOC1* repression to enable the plant to fully commit to flowering.

SEP3 induction also leads to other factors being recruited to begin to repress gene classes in certain whorls as shown in Figure 7.3 (O'maoileidigh *et al.,* 2014).

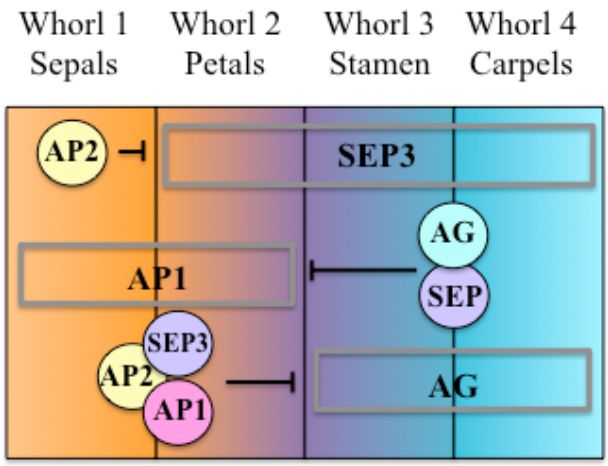

Figure 7.3. Patterns of gene induction and repression in the floral meristem that lead to the development of each distinct whorl of floral organs. Grey boxes represent promoters of the genes being repressed. The zone covered by the grey box indicates the area where expression occurs. Circles represent proteins being produced in whorls, which repress expression of other floral organ determination genes.

7.1.1.3 Proposed role of CDFs in the long day flowering pathway

In *Arabidopsis*, transgenic lines constitutively overexpressing *CDF* have been shown to delay flowering due to *CO* repression, whilst mutants of multiple *cdf*s have led to early flowering because of increased *CO* expression (Fornara *et al.,* 2009). CDF1 and 2 protein level is regulated by GI and FLAVIN-BINDING KELCH REPEAT F-BOX PROTEIN (FKF1). Blue light is required to stabilise the GI-FKF1 interaction, and so longer photoperiods lead to increased amounts of the complex, leading to increased degradation of CDF1 and CDF2, thereby mediating floral induction (Imaizumi *et al.,* 2005; Fornara *et al.,* 2009). ZTL also interacts with GI, and is able to target CDF2 for degradation in a blue-light dependent manner (Song *et al.,* 2014)

It has been shown these CDFs in *Arabidopsis* act redundantly as a mutation in one CDF can have little effect on flowering time. This redundancy has also been shown to be the case in tomato, which are day neutral plants (Corrales *et al.,* 2014).

In *Arabidopsis* the mRNA profile of three CDFs: 2, 3 and 5 were found to be high at the beginning of the light period and decrease during the day, with low level expression between 4h and 20h in 12:12 light/dark conditions (Fornara *et al.,* 2009). This was also the case in tomato, where *CDF1,2,3* and *5* all peaked at the very start of the light period (Corrales *et al.,* 2014).

CDF1 is the only DOF factor so far that has been shown to directly bind to the promoter of *CO*, and mediate its repression. As other CDF DOFs have been shown to be redundant with CDF1 through genetic approaches, it remains possible that other CDFs also bind to the *CO* promoter.

When *CDF2* was overexpressed in Arabidopsis, it caused a large reduction in *CO,* and also abolished *FT* expression (Fornara *et al.,* 2009). This further demonstrated that this family of transcription factors act to repress flowering. This supported the proposal that these DOF TFs play an antagonistic role to GI in modulating *CO* (Fornara *et al.,* 2009; Fornara *et al.,* 2015).

DOF factors have also been shown to be involved in red and far-red light signalling. For example, *COGWHEEL1* (*COG1*), is CDF-related TF which has been demonstrated to act as a negative regulator of both the phytochrome A and B (phyA and B) pathways (Park *et al.,* 2003). In Arabidopsis, it was shown that overexpression of *COG1* led to hyposensitivity to red and far-red light, whereas reduction in level led to hypersensitivity (Park *et al.,* 2003). OBF4 BINDING PROTEIN3 (OBP3) is another DOF that integrates both far-red signalling from PHYB, and blue light signalling by CRYPTOCHROME1 (CRY1), to influence hypocotyl elongation and cotyledon expansion respectively (Kang *et al.,* 2000).

7.1.1.4 Short day flowering

Whilst the majority of flowering time research has been focused on Arabidopsis which flowers in response to long days, more recently, researchers have used rice as a model for short day flowering plants.

From comparisons between the day-length sensing pathways operating in the two species, it appears that *FT* is highly conserved, but is regulated differently. In rice the homolog for *FT* is *HEADING DATE3 (Hd3)*, which is induced via the *CO* homolog *Hd1*. The mechanism for flower induction in rice therefore consists of Hd1 repressing *Hd3* in long days, but under short days, Hd1 induces *Hd3* expression, leading to flowering (Izawa *et al.,* 2002). *Hd1* and *CO* show the same expression pattern, and appear to both be regulated by the circadian clock. The rice GI homolog (OsGI) has also been shown to regulate *Hd1*, just as GI does with *CO* in *Arabidopsis* (Hayama *et al.,* 2003).

Research into flowering time has also demonstrated that whilst LDPs appear to find the length of the light period more important, SDPs are usually more dependent on a long dark period. SDPs are more sensitive to night break treatment, where impulses of light are introduced during the dark period, but are usually less sensitive to light quality during the light, in relation to flowering (Park *et al.,* 2003). Most LDPs are more sensitive to light quality and many require far-red light at the end of the long photoperiod in order for a long day to be interpreted as a signal leading to the induction of flowering.

Interestingly, flowering pathway components with no homologs in *Arabidopsis* have been identified in rice. One example is *EARLY HEADING DATE1* (*Ehd1*), which has been shown to promote flowering preferentially under short days, even without functional Hd1 (Doi *et al.,* 2004). Analysis suggests this gene functions upstream of

Hd3 (*FT*). It was shown to peak before dawn, and only displayed a high level of induction under short days (Wu *et al.,* 2008).

In other SDPs, such as the Japanese Morning Glory (*Pharbitis nil)*, it was found that the two FT homologs: *PnFT1* and *PnFT2* were only induced if the dark period was long enough (Hayama *et al.,* 2003). Furthermore in these plants depending on light regime *PnFT* and *PnCO* expression could be uncoupled, without effects on flowering. This therefore suggests that an additional pathway may exist for *FT* induction. This could mean SD plants have extra flowering induction pathways yet to be discovered.

7.2 Results

7.2.1 Impacts of changes in *KfCDF2* expression levels on the developmental progression into flowering

Despite *K. fedtschenkoi* being a SDP, *KfCDF2_FL* overexpressing transgenic lines were found to flower under 16:8 light: dark long-day conditions (LDs). The *KfCDF2* over-expresser lines were observed to transition to flowering at various times of the year, suggesting that the influence of *KfCDF2* over-expression on the induction of flowering was independent of the annual changes in day-length and light quality and quantity falling naturally on the greenhouse. Long-days (16:8) were maintained in the greenhouse using supplementary lighting, which maintained 16:8 conditions throughout the year.

It was therefore important to investigate this constitutive flowering phenotype in the *KfCDF2* over-expresser lines and to try to identify components of the molecular mechanism under-pinning this photoperiod independent flowering. These initial findings also suggested it was important to investigate both the over-expression and RNAi knockdown lines for *KfCDF2* under a variety of light regimes to gain further insights into the impact of *KfCDF2* levels on flowering and the speed of floral development.

Three pots, each containing three clonal plants of each line: wild type, *KfCDF2_FL_14C* or *KfCDF2_RNAi_19A,* were placed in three different Snijders Microclima MC-1000 growth cabinets under different light/ dark regimes. One cabinet was maintained under 16:8 LD (long day) conditions, which does not induce flowering in the wild type. A second cabintet was maintained under 12:12 LD, which can induce flowering if plants are maintained under LD conditions for several months,

and the third cabinet was maintained under 8:16 LD (short day) conditions, which induces flowering in *K. fedtschenkoi*.

Plants were grown for 80-days under these conditions. Figure 7.4 shows one pot of each plant type as they entered each condition, and Figure 7.5 shows the same pot at the end of the 80 days.

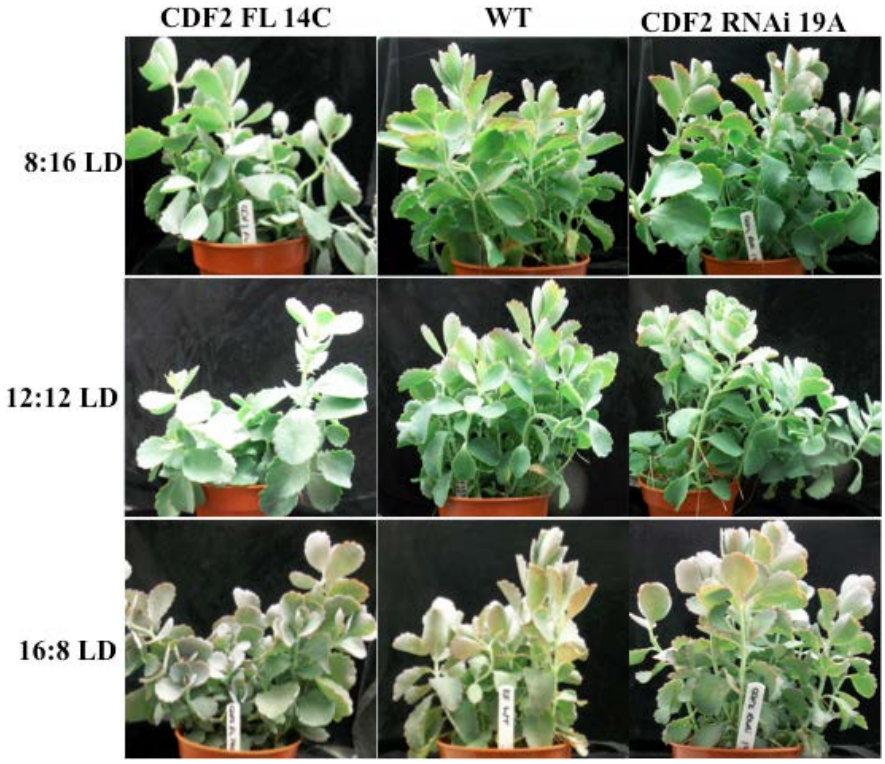

Figure 7.4. Starting phenotypes of *K. fedtschenkoi KfCDF2* transgenic lines subjected to different day-length regimes. *K. fedtschenkoi* wild type, *KfCDF2_FL_14C* and *KfCDF2_RNAi_19A* were grown in 16:8 LD (long day, non-flowering conditions) initially for 6 months. Plants were then placed into different growth cabinets under the different light regimes indicated on the left hand side of the figure.

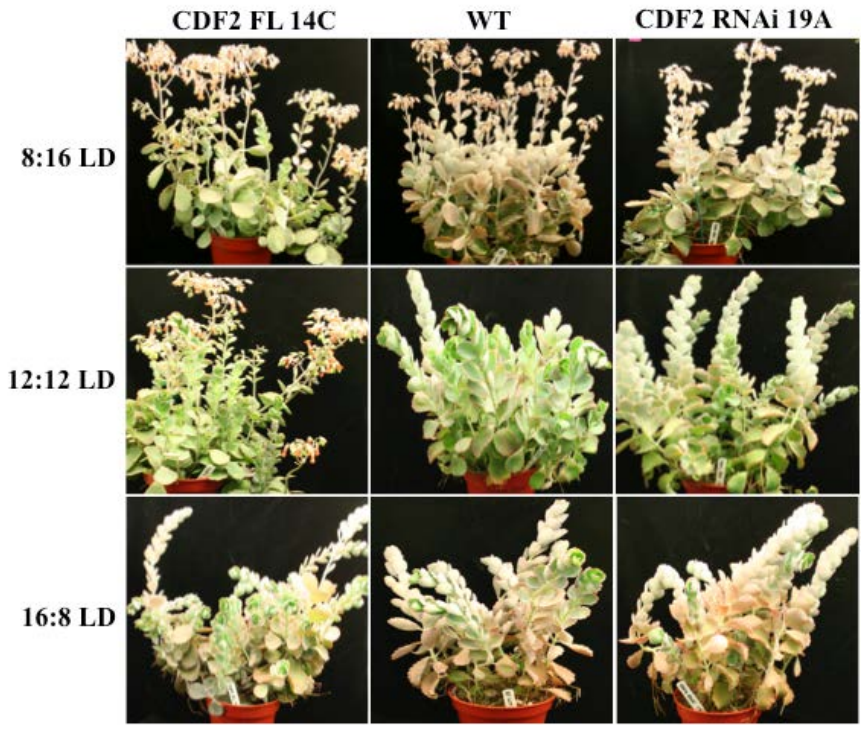

Figure 7.5. Progression of floral induction under different day-length regimes in wild type and *KfCDF2* over-expressor and RNAi liens of *K. fedtschenkoi*. *K. fedtschenkoi* wild type, *KfCDF2_FL_14C* and *KfCDF2_RNAi_19A* were grown under the indicated day-length regimes for 80-days in total. In 8:16 LD all three lines flowered, but *KfCDF2_FL_14C*s first flower opened 17-days earlier than the wild type and *KfCDF2_RNAi_19A*. In 12:12 LD conditions, *KfCDF2_FL_14C* also flowered only 3 days slower than in the 8:16 LD condition, whereas *KfCDF2_RNAi_19A* only showed a hooked stem, and wild type had yet to form a hook. In 16:8 no lines had hooked their meristematic region, but *KfCDF2_FL_14C* had started to lengthen its stems and was producing smaller leaves, which are both characteristic of the progression towared meristem hooking and floral initiation in *K. fedtschenkoi*.

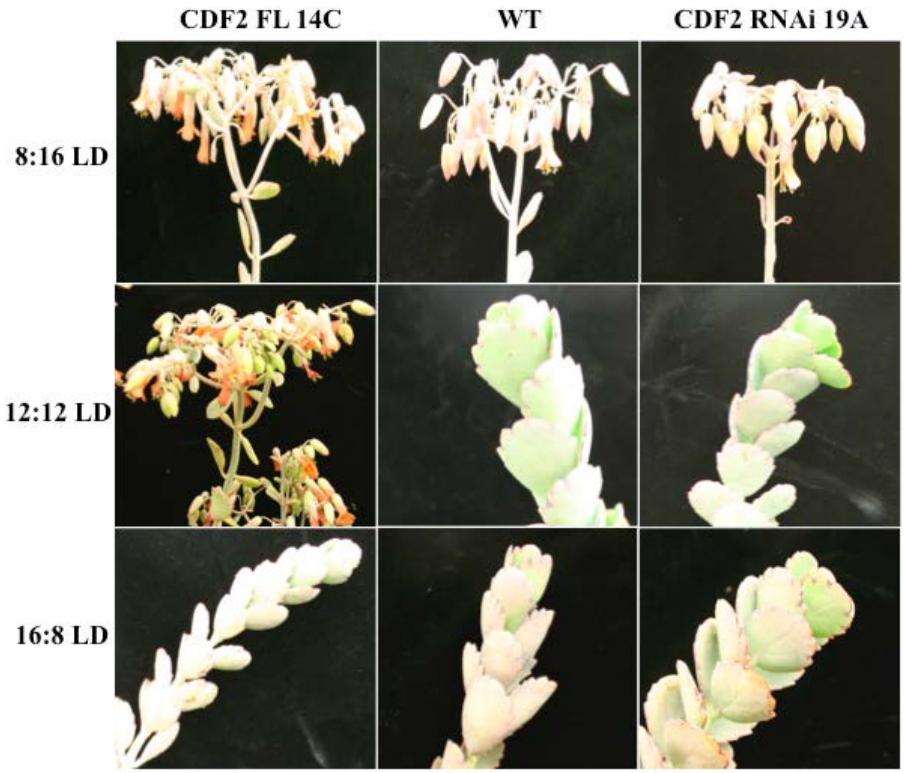

Figure 7.6. Close-up images of a stem from each of the *K. fedtschenkoi* lines at the end of 80-days in various light regimes. Light regimes are listed on the left-hand side and line names are listed above. By the end of the experiment, *KfCDF2_FL_14C* plant were producing many flowers in both 8:16 (short days) and 12:12 conditions. For Wild Type and *KfCDF2_RNAi_19A,* there was only one open flower, whilst the rest were still unopened buds under 8:16 SD conditions. For *KfCDF2_RNAi_19A* in 12:12 LD, the hooking-over of the mersitem could be seen. In 16:8 LD (long days) there was no visible hooking or flowering for any of the lines, but the stem of *KfCDF2_FL_14C* had elongated and the newly developed leaves on the elongated stem were smaller than those on the other two lines.

In 8:16 conditions (SDs), which induced flowering in all lines, the over-expressing line flowered 17-days earlier than the wild type and RNAi line. Flowering was classed as the day the first flower became visible and was fully open. Also, in short days, the *KfCDF2_FL_14C* plants had, on average, more open flowers (73) after the 80-day observation period, compared to wild type (3), or *KfCDF2_RNAi_19A* (2). Furthermore, in 12:12 LD there was a much more striking difference, as the wild types had yet to begin to hook after 80-days, whilst the *KfCDF2_FL* had produced 68 flowers (Fig. 7.5; Table 7.1). From the 12:12 results though, it appears that the RNAi

lines may also show a slightly more rapid flowering response than wild type due to visible stem hooking (Fig. 7.6). In long-days (16:8 LD), no hooking or flowering occurred, but it may take longer than 80-days to achieve the required threshold level of a floral induction factor in the *KfCDF2_FL_14C* line which had been observed to develop flowers under 16:8 LD conditions in the greenhouse. In the over-expressing line under the 16:8 long day conditions, stems had begun to elongate, and leaves had become smaller, characteristic of the start of flowering (Fig. 7.6). Furthermore, in the greenhouse, plants that had been kept under long days (16:8 light/dark conditions) for over 6-months were also not yet flowering, which may suggest that plant age plays an important role in flowering initiation in the *KfCDF2_FL* over-expresser lines.

Table 7.1. The influence of different day-length regimes on various traits indicative of the developmental progression towards flowering in *K. fedtschenkoi.*

	Days taken in 8:16	Days taken in12:12	Days taken in16:8
FL Hook	23	35	-
FL Flower	63	66	-
WT Hook	35	-	-
WT Flower	80	-	-
RNAi Hook	36	80	-
RNAi Flower	80	-	-

The number of days required to achieve stem hooking and the opening of the first flower under each light regime for each line tested. The experiment terminated after 80-days, and so dashes indicate that the line did not reach that flowering stage within the time of this experiment.

Under 12:12 light/ dark conditions the biggest differences in flowering induction was observed, but due to time constraints, and because all lines flowered under 8:16 light/ dark cycles, the gene expression investigation on the genetic control of flowering time was conducted using the samples collected from the 8:16 light/ dark conditions, so that the developmental progression towards flowering in all three lines could be observed and compared.

7.2.2 Gene expression changes in response to a flowering stimulus

To investigate gene expression changes involved in flowering time regulation in the *KfCDF2* lines, a time course was set up. Six-month old plants of each genotype were raised in non-inductive conditions (16:8 LD), and week 0 samples were collected. For all remaining weeks of the time course, the plants were under short days (8:16 LD). Wild type lines successfully flowered in 8:16 LD conditions after 84 days (12 weeks). Samples of a range of above-ground organs and tissues were collected each week to allow the investigation of flowering pathway gene transcript levels in various parts of the plant. The samples collected included LP1, 3, 6 and 10, stem and shoot apical meristem (SAM). As plants progressed to flowering, the hooked region of the stem at the growing point, the developing buds and flowers were also collected.

Various *K. fedtschenkoi* orthologs of flowering pathway genes were selected for investigation using RT-qPCR. The flowering genes were chosen based on their known roles in the regulation of flowering time in Arabidopsis. The Arabidopsis homologs were identified in the draft genome and assembled transcriptome of *K. fedtschenkoi,* and RT-qPCR primers were designed for their amplification and quantification. A gene referred to as *Contig9471* was used as the reference gene. This gene in *K. fedtschenkoi* has highest homology to At2g30720, which encodes a thioesterase found in the mitochondrion of cells. It was found to show very little change in expression level throughout the 24 h cycle, and so was used as the reference gene for the relative quantification of gene transcript abundance levels for the flowering pathway genes.

Table 7.2. Identified flowering genes in *K. fedtschenkoi* with accession numbers in Arabidopsis and Chrysanthemum.

Gene Name	Symbol	*Arabidopsis* Accession Number	Number of hits in *K.f.*	*C. seticuspe f. boreale* Accession Number	*Kf* Genome accession numbers	*Kf* gene symbol
CYCLING DOF FACTOR2	*CDF2*	At5g39660	1		KF17950	*KfCDF2*
GIGANTEA	*GI*	At1g22770	1		KF96760	*KfGI*
CONSTANS	*CO*	At5g15840	2		KF105990	*KfCO25*
					KF127450	*KfCO46*
ZEITLUPE	*ZTL*	At5g57360	1		KF118840	*KfZTL*
SUPRESSOR OF OVEREXPRESSION OF CONSTANS1	*SOC1*	At2g45660	1		KF03655	*KfSOC1*
SEPALLATA	*SEP2*	At3g02310	1		KF22770	*KfSEP*
	SEP3	At1g24260				
FRUITFUL	*FUL*	At5g60910	1	AB839766	KF111895	*KFFUL*
APETALA1	*AP1*	At1g69120	1	AB839766	KF135520	*KfAP1*
FD	*FD*	At4g35900	1	AB839768 /9	KF106650	*KfFD*
FLOWERING LOCUS T	*FT*	At5g03840 At1g65480	2	AB839766 AB839767	KF67365 ----	*KfFT12* *KfFT18*

Genes involved in flowering time that were identified in Arabidopsis (TAIR entries) and *Chrysanthemum seticuspe f. boreale* (DDBJ entries)(Higuchi *et al.,* 2013) and orthologs that were indentified within the *K. fedtschenkoi* transcriptome and genome. *KfFT18* could not be found in the draft assembly of the *K. fedtschenkoi* genome, but was found in the assembled transcriptome as Contig18180.

7.2.2.1 Wild type progression to flowering

Initial samples were collected on day 0 from plants that had been grown under long-day (16:8 LD) conditions since they were small, developmentally-synchronised adventitious leaf-margin plantlets. The light regime was changed to short days (8:16 LD). Sampling was then conducted after various numbers of weeks: 1, 5, 7, 9, 11 and 12. After 5-weeks, the plants had developed a noticeably hooked over stem at their growing point. By week 7, small buds had begun to form, and by week 11 large buds were forming. In the 12th week, flowering occurred in wild type.

In older leaves, *KfCDF2* was an abundant transcript at week 0, consistent with its original identification as putative CAM-associated, clock-controlled TF (see Chapter 3). However, in all other tissues, *KfCDF2* transcripts were low until after week 5 of SD conditions (Fig. 7.7). *KfCDF2* was thereafter present in all tissues at a high transcript abundance until buds began to develop, at which point *KfCDF2* transcript levels declined. In the meristems, dark levels of *KfCDF2* increased between week 5, when the stems began to hook, and week 11, when bud development had been established. These results for wild type plants suggested that *KfCDF2* plays a role in promoting the developmental progression to flowering, but is of lesser importance during floral development in the meristematic tissues.

As *KfGI* is a core circadian clock gene, it is not surprising that its transcript levels remained relatively constant in the leaves throughout the short-day inductive timecourse (Fig. 7.8). After 7-days under short days, *KfGI* increased in transcript abundance in the light in all leaf pairs (week 1), which is likely sending an inductive signal to the flowering pathway. *KfGI* transcript levels subsequently began to increase in the light in the stem, up to week 9, when visible buds had developed. Levels of *KfGI* in the meristem itself did not change dramatically over the duration of the experiment, although the light peak of *KfGI* did show a general decline as the weeks under short day conditions progressed. During bud development, the level of *KfGI* increased, but subsequently dropped when the flowers opened, suggesting *KfGI* may play a role in regulating floral development, or have some role in floral development itself, in addition to its role in linking the clock and flowering time pathways.

For *CO*, two partial sequences were found in the *K. fedtschenkoi* transcriptome, which show high sequence homology: *KfCO25* and *KfCO46*, and so both transcripts were investigated (Fig 7.9 & Fig. 7.10). Their nucleotide sequences show 84.5 % identical bases, and at protein level, their predicted amino acid sequences shared 83.5% pairwise identify. Both genes showed very similar results, except that *KfCO46* was a lower abundance transcript, and also appeared to show a less robust oscillation between the light and the dark sample.

Both *KfCO* genes were not detected at either the light or dark timepoints samples in this experiment before week 5. In week 5, both *KfCO* transcripts were detected in leaves, and their transcript abundance then increased steadily increases. In leaves and the meristem, *KfCO* transcript abundance peaked in week 11, whilst in the stem it peaked at week 12. *KfCO25* transcript levels declined steadily in buds as they developed to flowers, whereas *KfCO46* levels remained high, in both light and dark, until week 12, when flowers opened. Based on these results, *KfCO46* could play a role in floral determination, whereas *KfCO25* may play a larger role in the progression from the vegetative meristem to the inflorescence meristem.

For FT, two genes were detected in the *K. fedtschenkoi* genome: *KfFT12* and *KfFT18*. Amino acid sequence alignments indicated that *KfFT12* shared highest sequence homology with antiflorigen (At5g03840), whereas *KfFT18* shared greatest homology with florigen (At1g65480). However, the RT-qPCR results presented here indicate that *KfFT18* may act as the antiflorigen in *K. fedtschenkoi* (Fig. 7.11 & Fig. 7.12). *KfFT18* transcript levels were found to be high in non-flowering conditions, after which levels began to decrease, especially in the light period, where this gene had displayed higher transcript levels relative to the dark. *KfFT18* was barely detected in

buds or flowers, except at week 12, when it was detected primarily in the dark period. *KfFT12* on the other hand had low transcript levels during non-inductive flowering conditions, and levels began to increase steadily once the plants were placed into conditions that promoted flowering (short days). *KfFT12* also peaked during the dark rather than the light, and was found to be present at very high transcript levels in buds and flowers. FT is thought to be a key component of the mobile flowering integrator signal, which is able to travel from leaves into the stem up to the meristem to mediate the induction of flowering.

For *KfAP1*, which is important not only for progression to flowering, but also in floral development, the transcript levels were high throughout most of the samples and weeks (Fig. 7.13). In LP10 once entering into flowering conditions, expression levels increased each week. In the meristem, levels of *KfAP1* were higher from week 7 onwards, and transcript levels were also shown to increase in buds as they develop, peaking in well-developed buds and flowers at week 12. Interestingly, in the stem there appeared to be a large final signal sent via the stem at the end of week 12, coincident with the flowers opening, which may suggest a role for this gene in the induction of flower opening.

With *KfSEP,* it was detected in all tissues at all times, except for the light period of week 0 (in long days; Fig. 7.14). In all tissues there appeared to be repression in light in non-flowering conditions (long days). Once plants entered short days, this repression seemed to be removed (Fig. 7.14). In some weeks, *KfSEP* transcript levels were higher in the light, whilst in other weeks transcript levels were higher in the dark, suggesting this gene may not be under circadian control in *K. fedtschenkoi*, but

this may be a result of phase changes in the timing of peak transcript abundance which could only be characterised with a more detailed timecourse with samples collected at more regular intervals over the light/ dark cycle. From all the tissues and timepoints measured, *KfSEP* transcript levels were at their highest in the meristems and buds, especially when the buds were large and beginning to open (week 12). These findings are consistent with the known role of *SEP* in co-regulating floral development all flowering whorls, which suggests it may also be involved in all stages of floral development.

KfZTL is another gene that functions within the clock, where it has specifically been demonstrated to function in targeting TOC1 ubiquitination and degradation via the 26S proteasome. ZTL has also been shown in Arabidopsis to mediate the degradation of CDF protein, especially CDF2 (Fornara *et al.,* 2009). Near the end of this flowering experiment, after buds became visible, *KfZTL* levels increased in all tissues (Week 11), with the highest transcript levels detected in LP10 (Fig. 7.15). *KfZTL* also reached its peak transcript levels in floral organs when buds were beginning to open, suggesting that it could be involved in targeting other flowering pathway proteins for degradation, to enable flowers to open and prevent floral reversion.

KfSOC1 is induced to promote flowering. In older leaf pairs, (LP6 and 10) there were high levels of the transcript in the dark period in non-flowering conditions (Fig. 7.16). As progression to flowering continued, *KfSOC1* transcript levels were detected in LP1 and 3, whilst lower levels were detected in LP6 and 10, likely due to it being repressed to prevent floral reversion. Levels also began to decrease in the meristem and the stem between weeks 7 – 9, and this may possibly also be a sign that this

regulation is linked to the plant avoiding floral reversion. Furthermore, *KfSOC1* levels increased steadily in buds, and peaked at bud opening, and so *KfSOC1* may play a further role in the final flower-opening signal.

FD is thought to interact with FT and induce flowering genes. Interestingly, in *K. fedtschenkoi* the transcript levels of *KfFD* were seen to decrease once plants entered inductive conditions for flowering. *KfFD* levels began subsequently to rise again from week 6 in older leaves, but were still significantly lower than in Week 0 (Fig. 7.17). In the stem and meristem, levels of *KfFD* were high until Week 5 and then declined. *KfFD* also increased in floral organs throughout development. This may be because if *KfFT12* is moving to the meristem, then this is the only place that *KfFD* transcript levels need to be maintained at a high level to enable the FT-FD complex to form.

KfFUL, which is essential for flowering, was barely detectable in any tissue in non-flowering conditions, but once the conditions had changed to 8:16 LD, levels of this transcript began to increase until the approximate time of stem hooking (week 5; Fig. 7.18). This supports the proposal that this gene maintains a conserved function between Arabidopsis and *K. fedtschenkoi*, being involved in the transition from inflorescence to floral meristem. *KfFUL* levels also increased in buds during their development, with slightly lower levels found in flowers.

Finally, *LEAFY* (*LFY*), which is another essential gene in the pathway of flowering time regulation and floral development in Arabidopsis (Fig. 7.1), was only able to be identified in the *K. fedtchenkoi* genome (KF07545). It was also amplified successfully from genomic DNA of *K. fedtschenkoi* using the designed qPCR primers. However,

KfLFY could not be found though in the assembled transcriptome data, nor could it be amplified in pooled flowering time samples using RT-qPCR. *KfLFY*s role in flowering progression could not therefore be determined in this study, as its expression could not be confirmed, suggesting LFY may not play an active role in *K. fedtschenkoi* development, or that its expression was missed in the range of samples used in this experiment; for example its expression may be confined to a particular time of day or to specific tissues or cell types which were not well represented in the RNA samples used in these experiments.

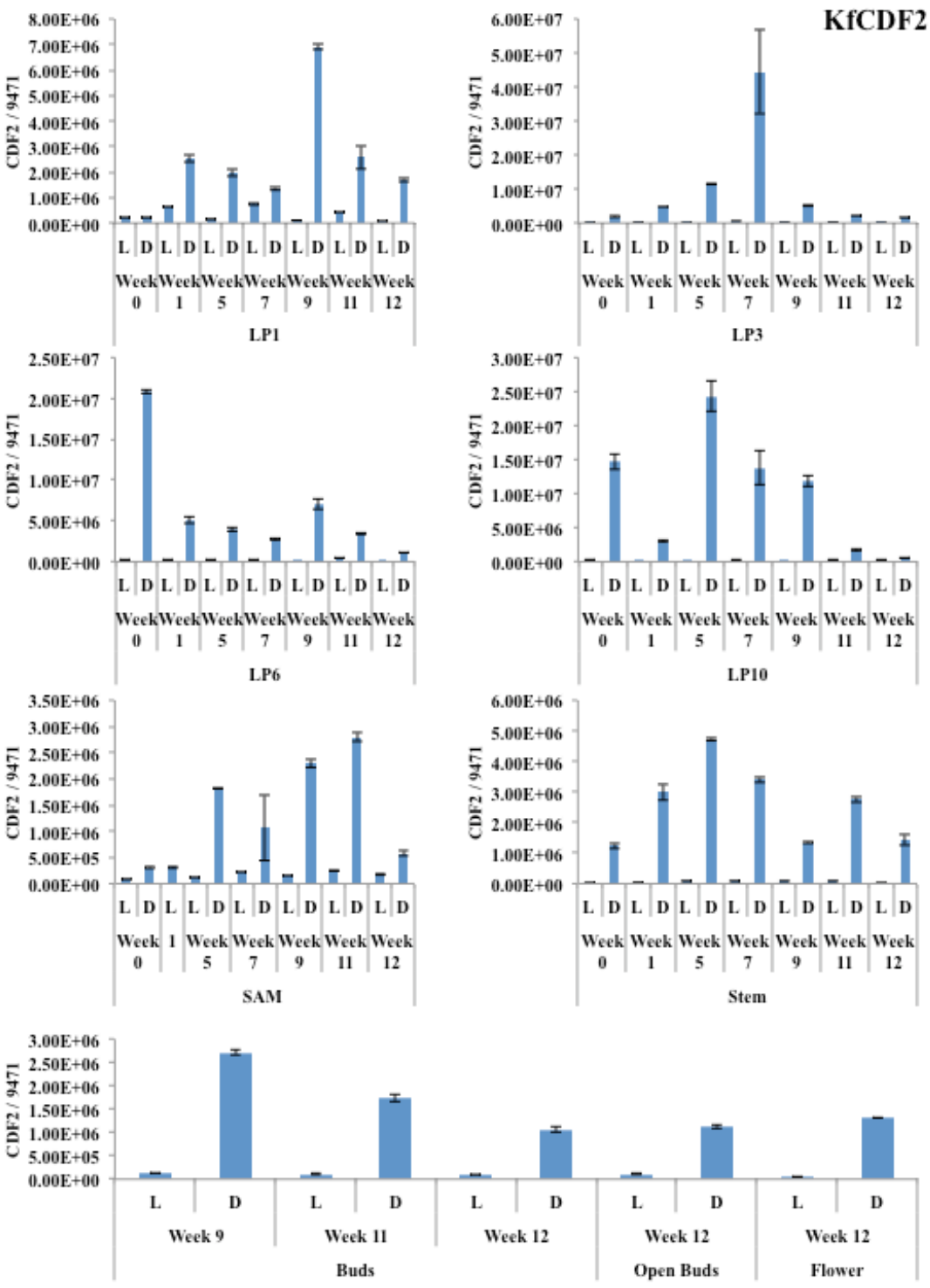

Figure 7.7. RT-qPCR data for *KfCDF2* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

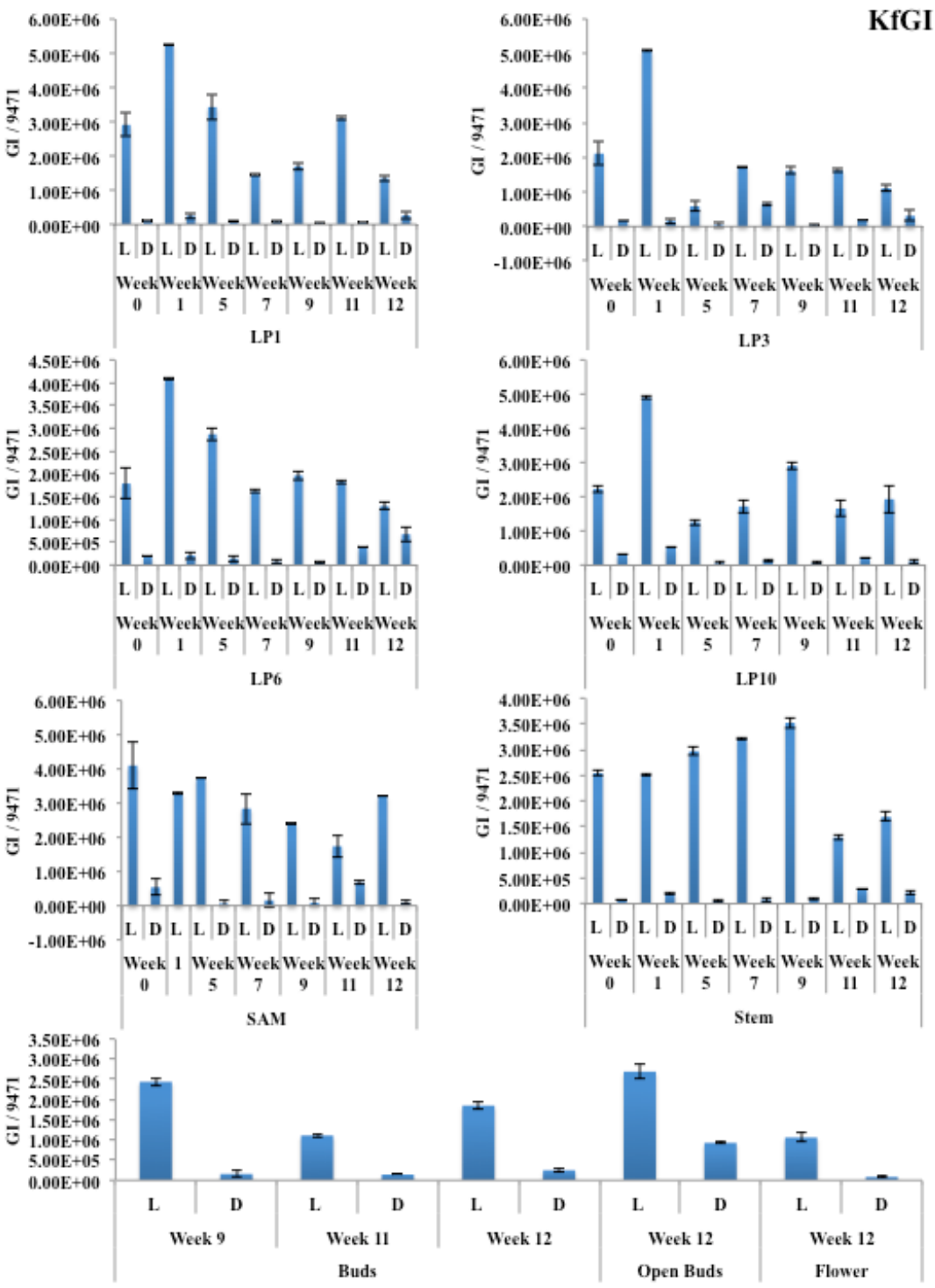

Figure 7.8. RT-qPCR data for *KfGI* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

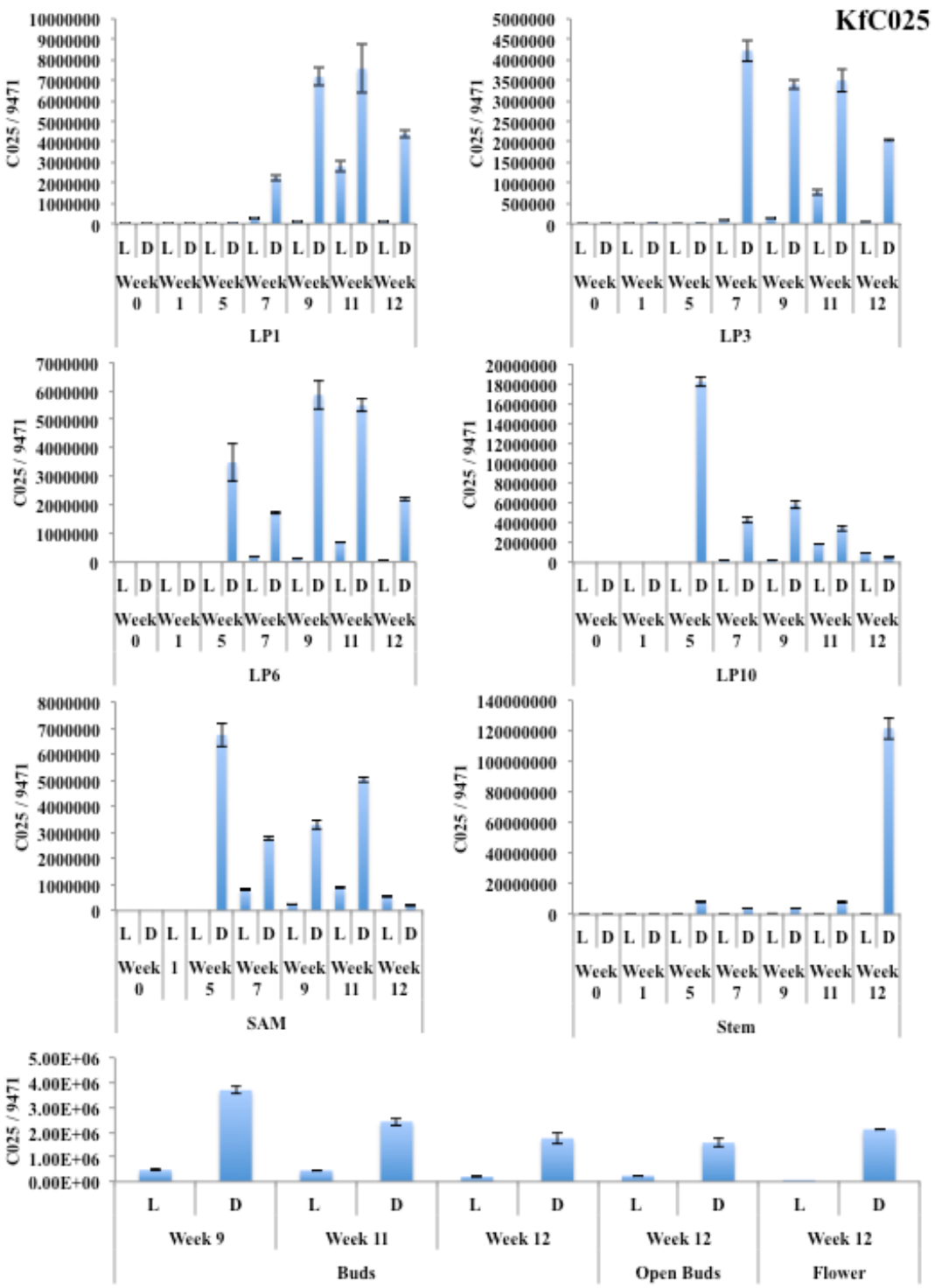

Figure 7.9. RT-qPCR data for *KfCO25* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

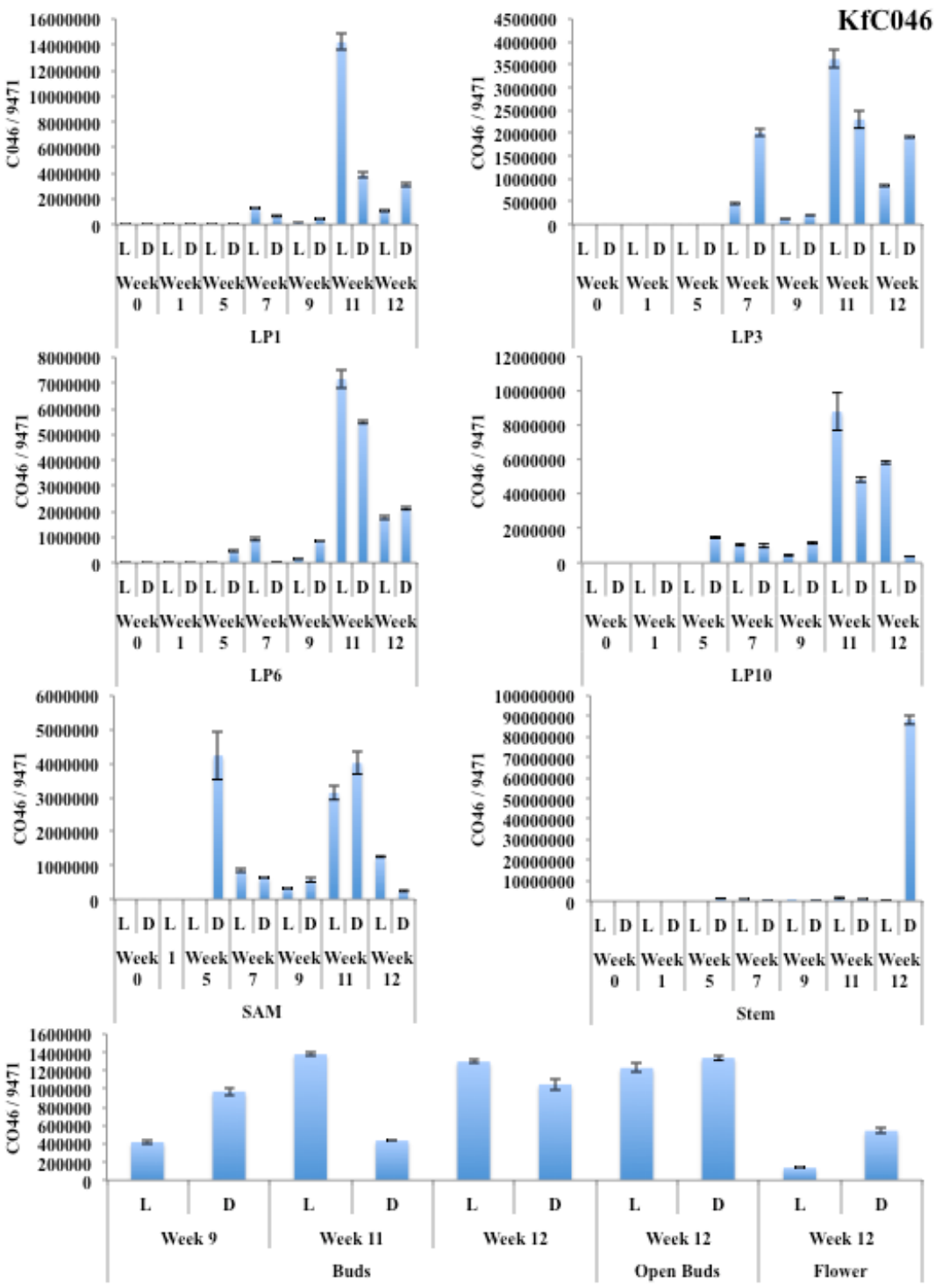

Figure 7.10. RT-qPCR data for *KfCO46* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

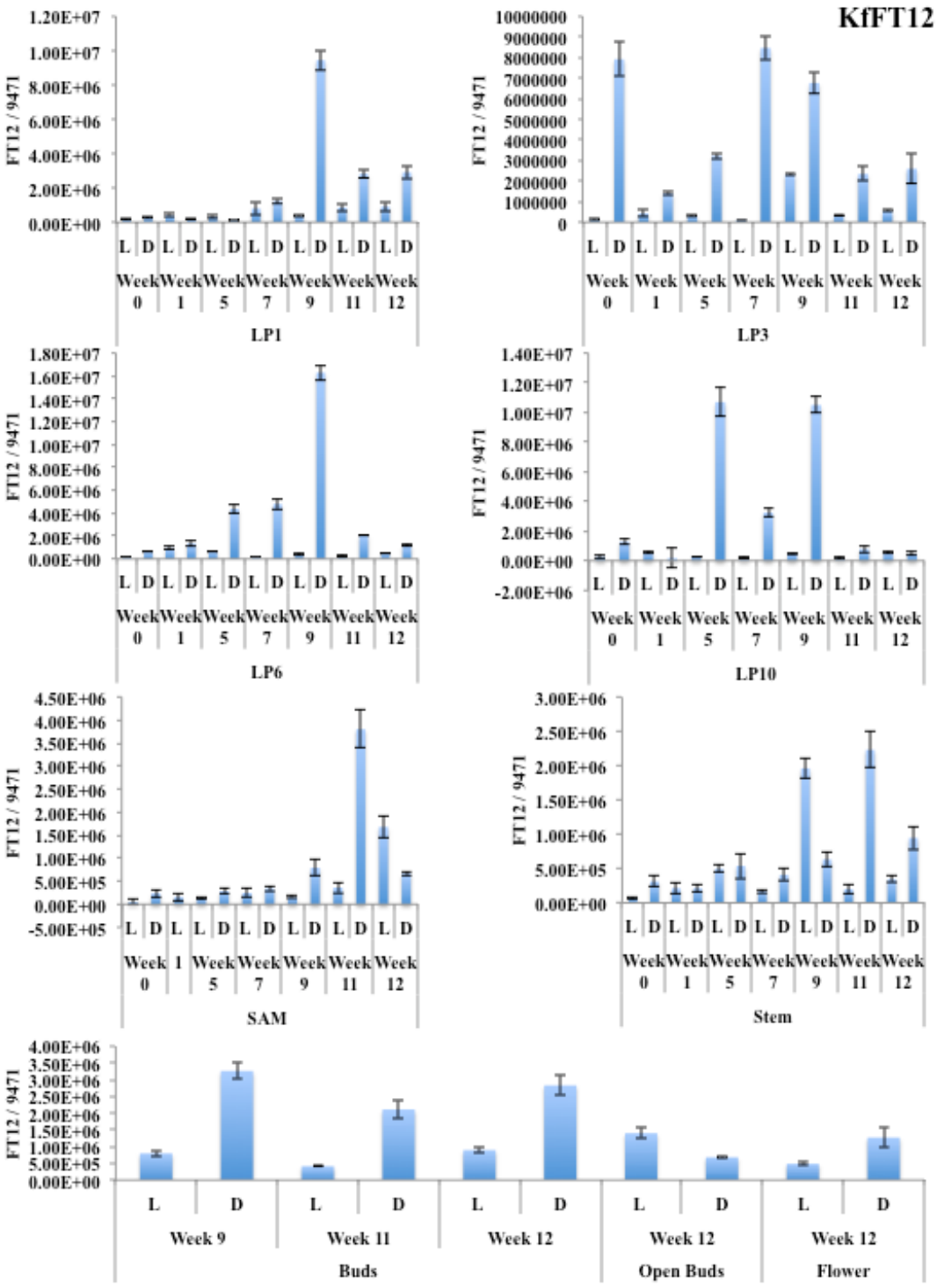

Figure 7.11. RT-qPCR data for *KfFT12* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

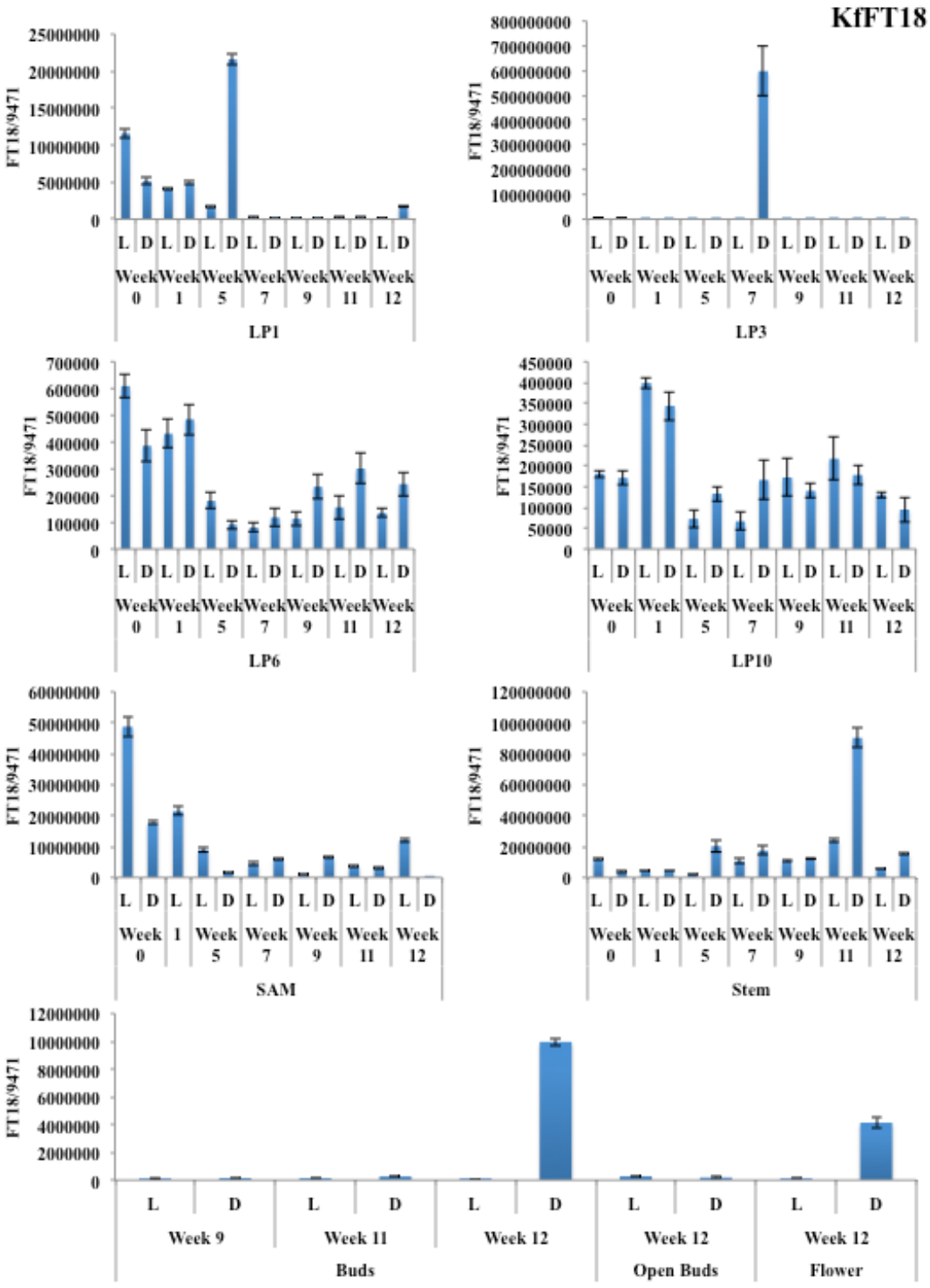

Figure 7.12. RT-qPCR data for *KfFT18* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

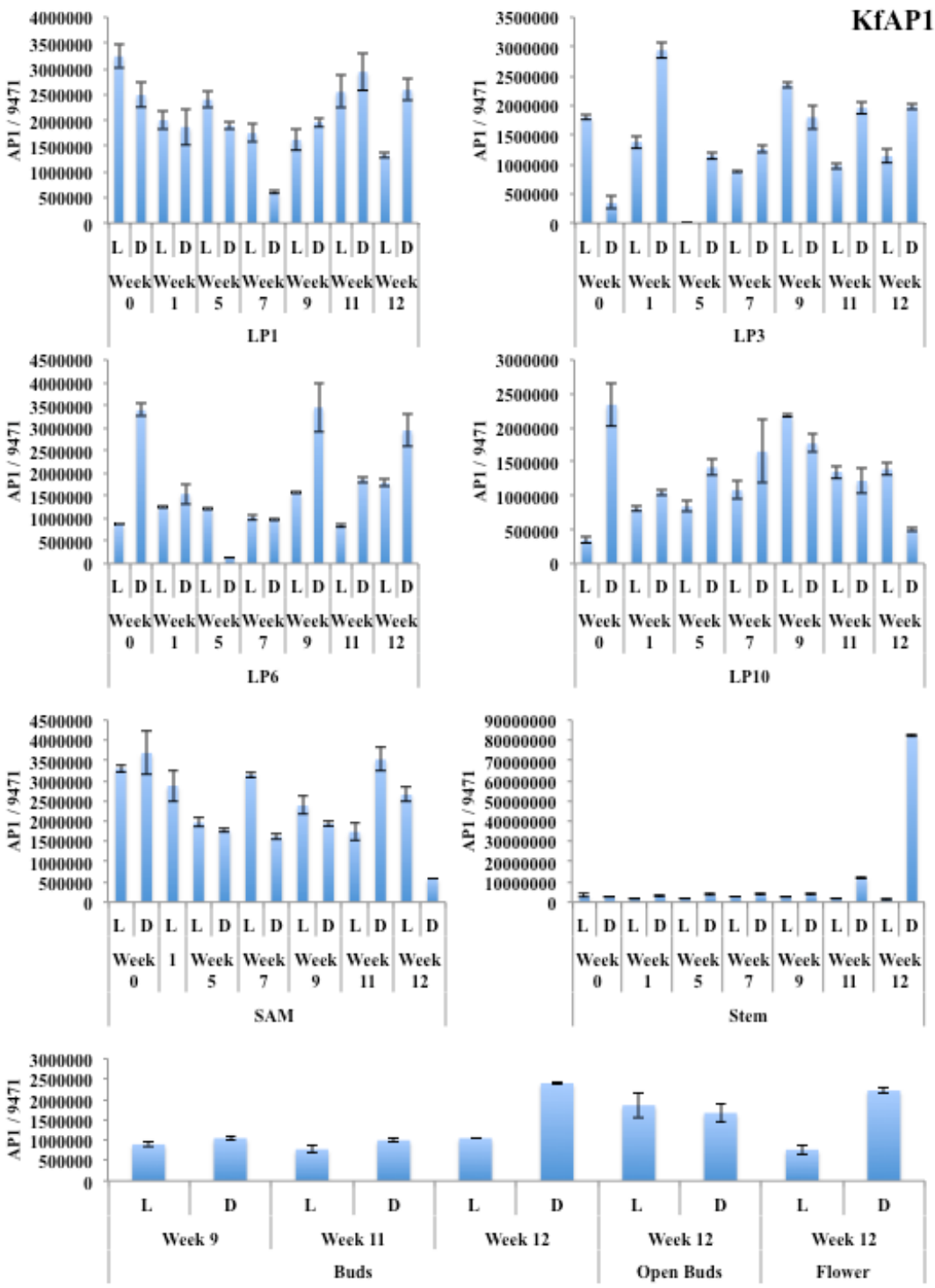

Figure 7.13. RT-qPCR data for *KfAP1* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

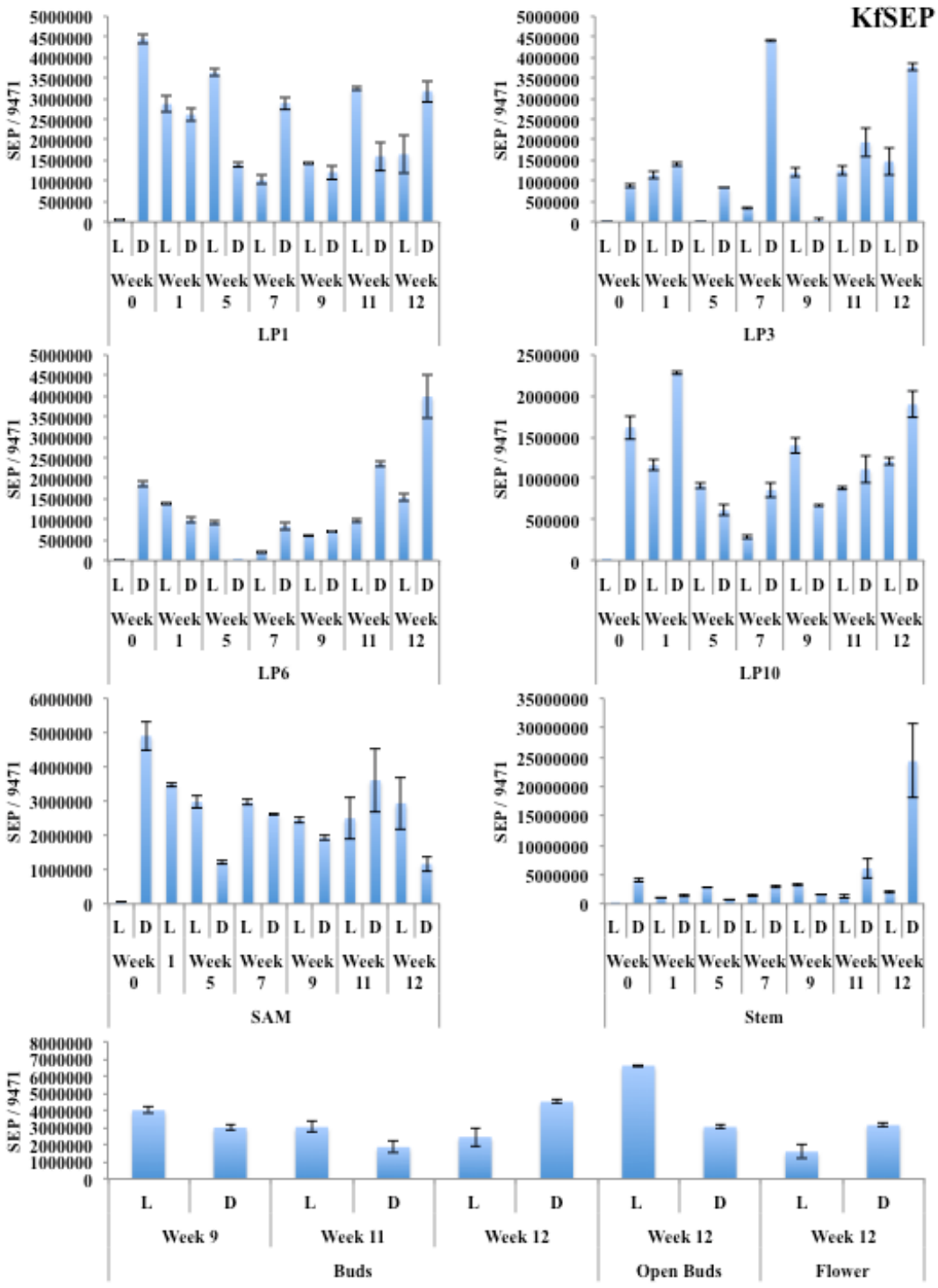

Figure 7.14. RT-qPCR data for *KfSEP* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

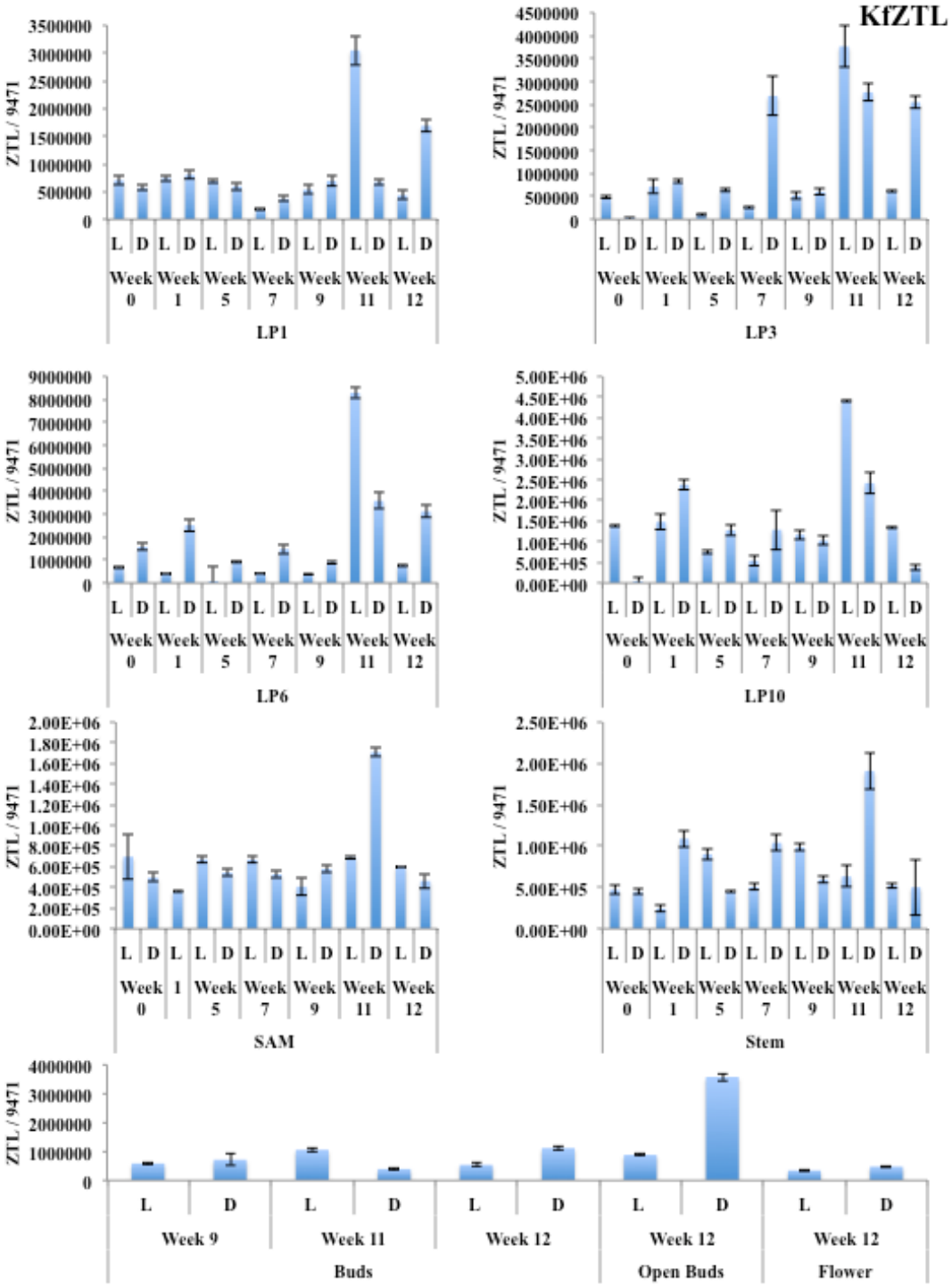

Figure 7.15. RT-qPCR data for *KfZTL* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

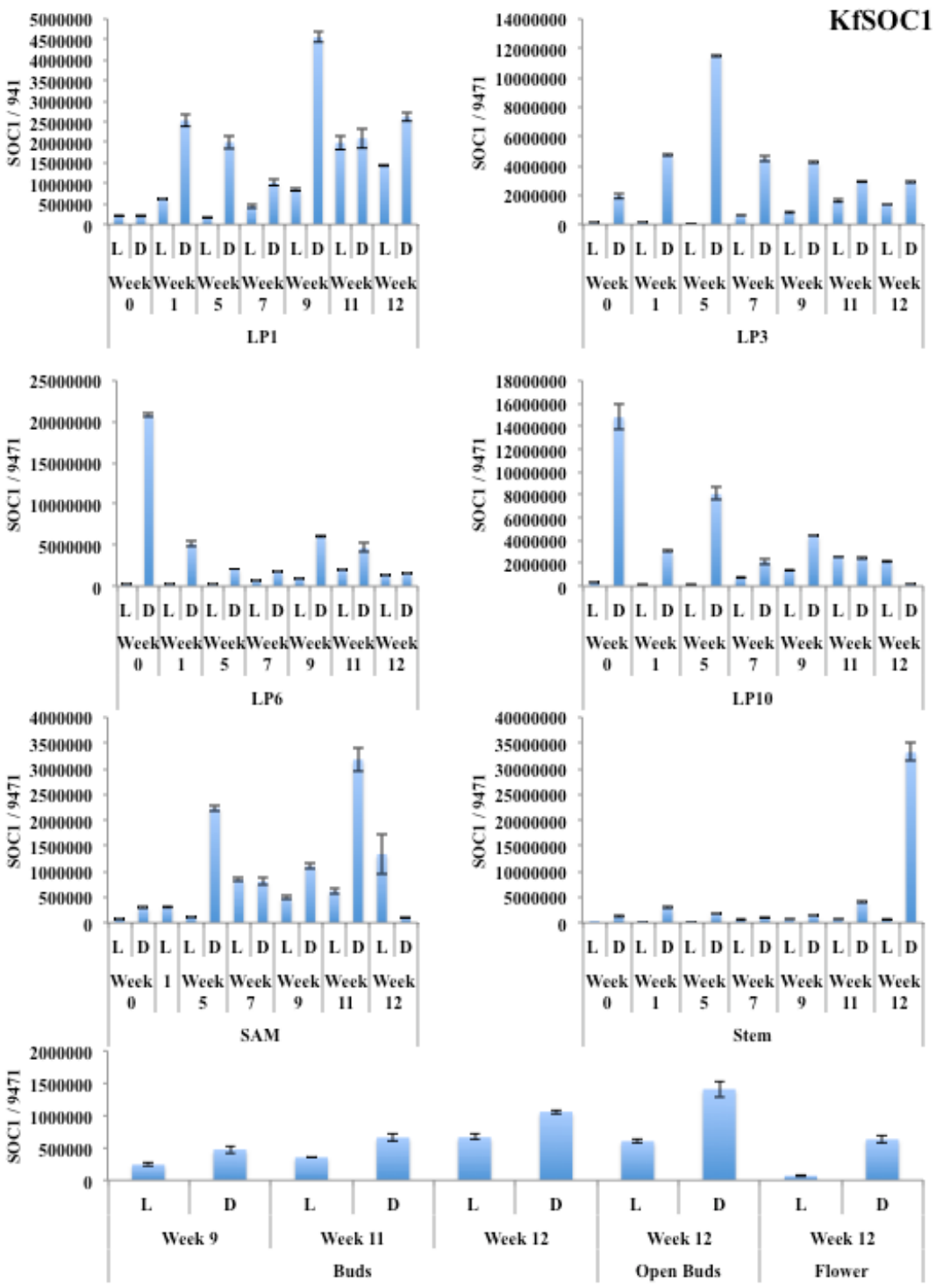

Figure 7.16 RT-qPCR data for *KfSOC1* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

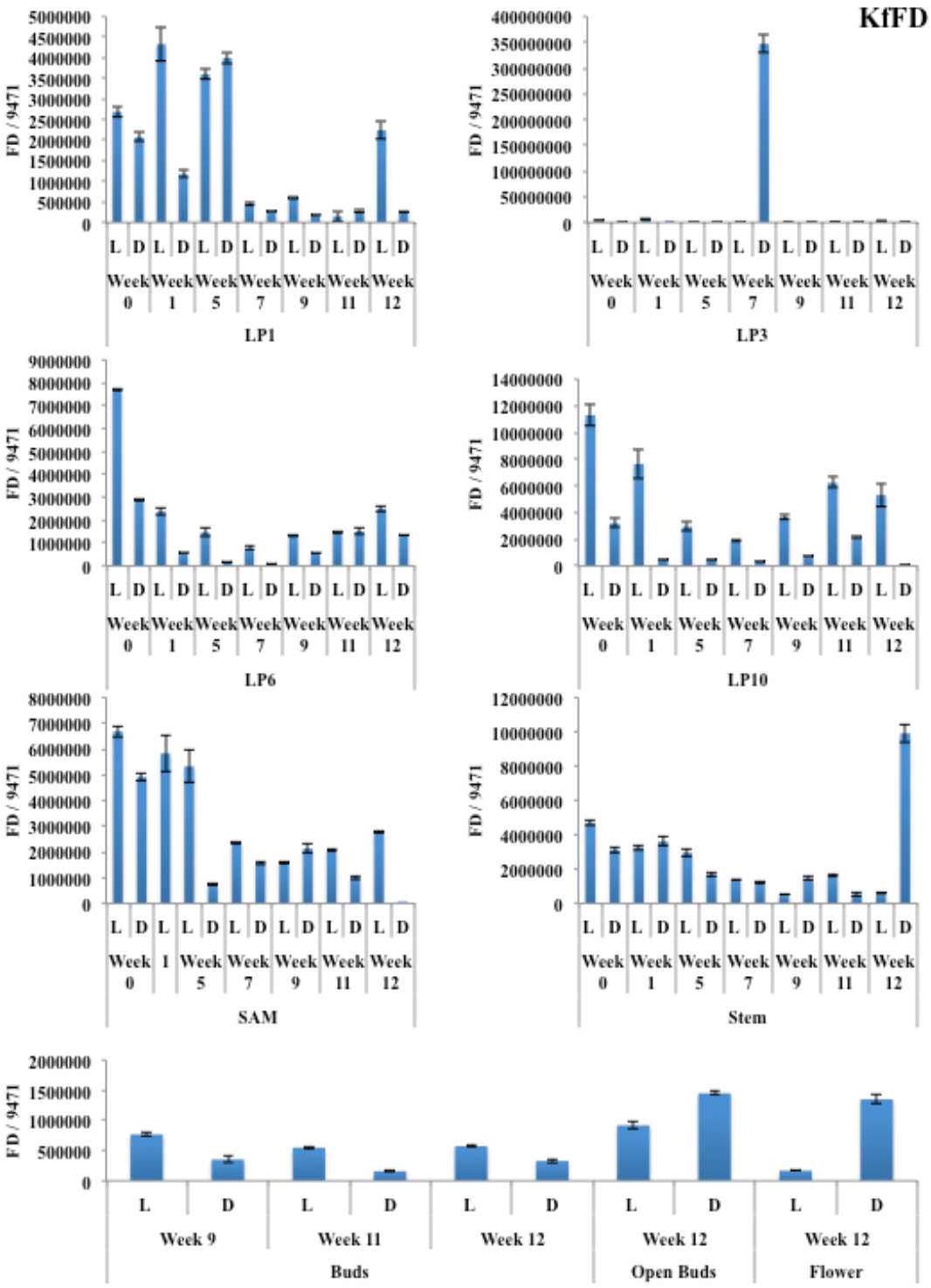

Figure 7.17. RT-qPCR data for *KfFD* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

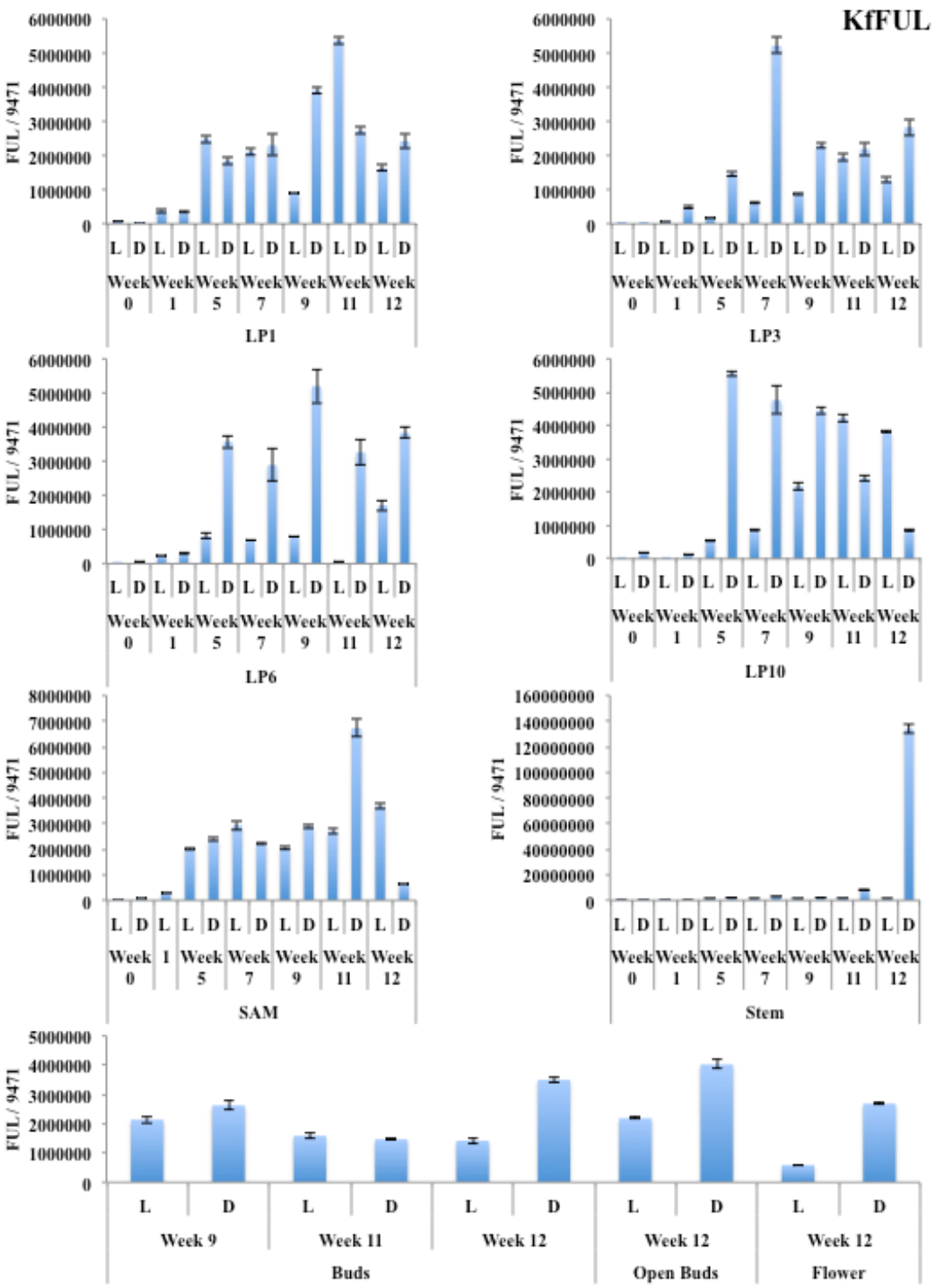

Figure 7.18. RT-qPCR data for *KfFUL* transcript levels in different tissues during the short-day induction of flowering in *K. fedtschenkoi*. Samples are: leaf pairs: 1, 3, 6 and 10, shoot apical meristem, stem and buds and flowers. L and D represent light and dark samples taken in that week. RNA extraction failed for the dark sample of the meristem in week 1. Normalising gene 9471 was used (Kf gene: KF149435).

7.2.3 Changes of the transcript levels of flowering pathway genes in *KfCDF2* transgenic lines under non-flowering conditions

From observations using plants kept in the greenhouse under long-day conditions (16:8 LD), and also from the initial flowering time experiment (7.2.1), it had been demonstrated that the *KfCDF2_FL* over-expresser lines were able to initiate flowering, and also flower more rapidly, than wild type in all conditions. Also, in 12:12 LD conditions, as shown in Figures 7.5 and 7.6, the RNAi line developed apical hooks, whilst the wild type had yet to do so, suggesting these lines may also be able to develop flowers more rapidly under certain day-length conditions.

To investigate the regulation of genes hypothesised to function in the photoperiodic flowering time pathway, transgenic lines were grown for 6 months in 16:8 conditions in the greenhouse, and then placed in a large walk-in growth room under short-day conditions; where wild type plants had been shown previously to flower successfully. Wild Types used in this experiment also managed to produce flowers after 12 weeks. Interestingly though, these conditions prevented the *KfCDF2* transgenic lines from flowering, or even developing any signs that they were attempting to initiate flowering. The light spectrum between the greenhouse, where *KfCDF2_FL* lines flower in long days, the Snijders growth cabinets used for the initial experiment (7.2.1), and the walk-in growth chamber were therefore compared to assess the quality and quantity of light in each growth area (Fig 7.19).

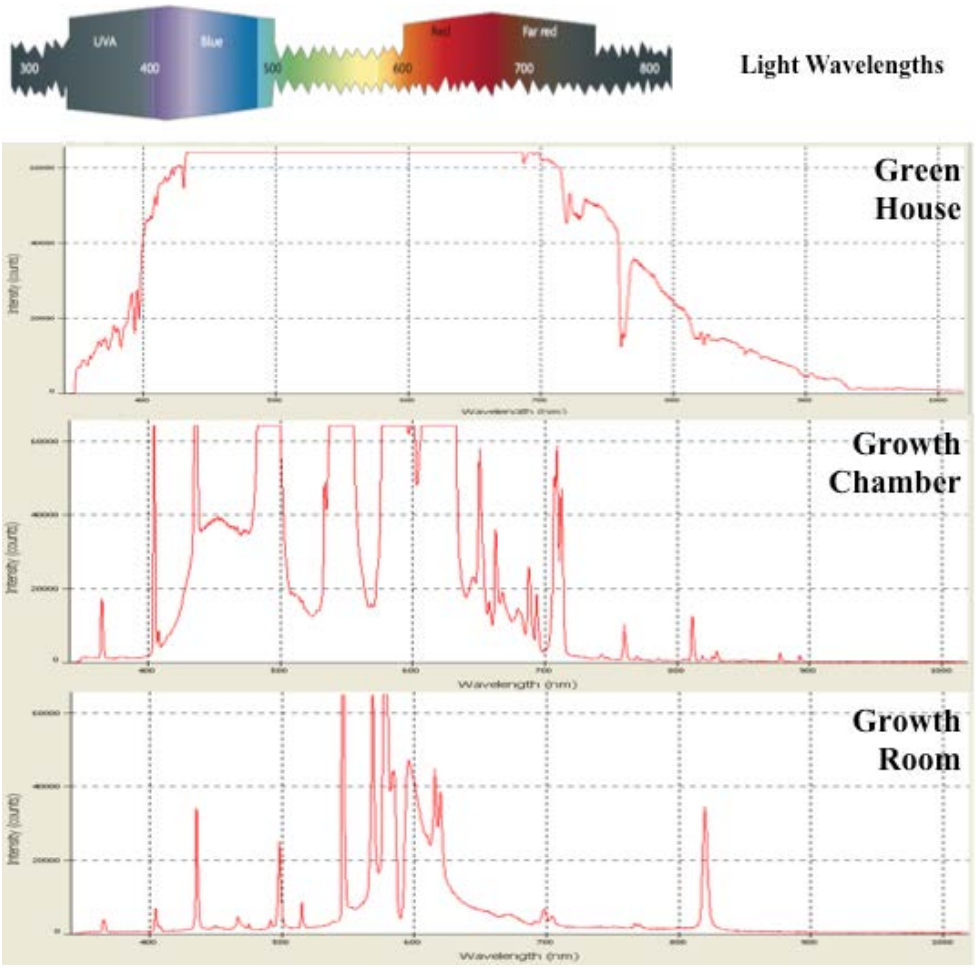

Figure 7.19. Light spectrum (quality of light) found in different growth areas. Greenhouse measurements were taken on an overcast day, during the morning and afternoon, but no difference in readings was detected. The morning reading is presented. Both greenhouse and growth chamber produced a better quality and quantity (intensity) of light than the growth room. Despite this, wild type plants still flowered in the growth room, suggesting that the *KfCDF2* transgenic lines were most susceptible to flowering inhibition when receiving either low quality and/ or quantity of light. Light Spectrum image taken from Jiao et al., (2007).

The light spectrum was compared within the three areas: greenhouse, Snijders Microclima MC-1000 growth chamber, and walk-in growth room containing high intensity white lights and also sodium lamps. As the growth room was the only area where *KfCDF2* transgenic lines did not flower, this spectrum was compared to that of the other two areas. The growth room had a lower quality of light compared to the other two areas (Fig. 7.19), as there were relatively low amounts of blue, red and far

red light emitted by the lamps in the growth room (Fig. 7.19; growth room). These wavelengths of light have been demonstrated in other species to be essential for the day-length dependent induction of flowering. It is thus noteworthy that wild type plants were able to flower under the growth room conditions. This suggests that changes in the transcript level of *KfCDF2* affected the plant's sensitivity to either blue and/ or red/far-red light. It therefore appears light quality is very important for *KfCDF2_FL* flowering. Furthermore, the quantity of light across most of the visible part of the spectrum was also lower in the growth room conditions, which could indicate that the *KfCDF2* transgenic lines required a high quantity of light in order to induce flowering.

Based on these findings, RT-qPCR was used to investigate the regulation of flowering pathway genes in the *KfCDF2* transgenic lines, *KfCDF2_FL_14C* and *KfCDF2_RNAi_19A* as well as the wild type in long days (16:8 LD) (Table 7.3 & Table. 7.4). All RT-qPCR graphs are presented in the appendix (S7.1 to S7.12) providing the full details of the magnitude of changes in transcript levels. To simplify the interrogation of this data, summary tables are presented here to highlight key trends within the gene transcript abundance data. In Tables 7.3 and 7.4, green squares represent up regulation compared to the wild type, red indicated down regulation and purple shown no significant change to the wild type.

7.2.3.1 Changes in the transcript abundance of flowering pathway genes in *KfCDF2_FL_14C*

In the *KfCDF2_FL_14C* line, *KfCDF2* was over-expressed in various tissues, with the only exceptions being LP3 and the dark period in the stem. This suggests that there

may be changes to *KfCDF2* regulation in different tissue types, depending on their importance in that specific area.

In *KfCDF2_FL_14C*, there were clearly genes that were up-regulated throughout the day and also throughout tissues, namely *KfCO25, KfCO46, KfAP1* and *KfFUL*, which are all genes that promote floral development, and must be up-regulated for flowering to occur. This could therefore explain why the *KfCDF2_FL* overexpressing plants flowered early, as they possessed elevated levels of these transcripts, which signalled that the external conditions are correct for flowering. As soon as these *KfCDF2_FL* plants were old enough, and thus had sufficient resources, it appears that flowering was initiated even in long day conditions. Other genes, such as *KfFT12, KfZTL* and *KfSOC1* also showed increased transcript levels at most times of day and in all tissues (Table 7.3). *KfFT18*, which from the wild type RT-qPCR data tracking the progression to flowering appeared to be involved in the repression of flowering, was also down-regulated in many leaf pairs/ organs, including LP1, LP3 and LP6, and the meristem in the light. This may explain why *KfFT12* and *KfAP1* were elevated in many of these leaf pairs/ organs, as levels of the putative repressor had been lowered. *KfSEP*, which is involved in controlling floral organ development in the different whorls of the developing floral meristem, was found to be under tight control. In the light period sampled, *KfSEP* transcript levels were elevated in all leaf pairs and organs, whereas in the dark sample this gene was down regulated. This could explain why flowering occurred more rapidly in these transgenic lines as a key gene involved in the regulation of the various whorls of floral organs was up-regulated, possibly leading to faster rates of floral organ development.

FD in Arabidopsis has been shown to interact with FT to induce many flowering genes, including *AP1*. It is therefore interesting that in this experiment in *K.*

fedtschenkoi, the *KfCDF2_FL* lines showed reductions in *KfFD* in half of the samples. However, based on understanding of the role and function of FD in Arabidopsis, FD is likely to interact with FT mainly in the stem and meristem, both of which only showed only very slight reductions compared to the wild type. Thus, these detected changes in *KfFD* levels may not impact on the progression to flowering in *K. fedtschenkoi*.

Two clock genes also had their expression profile analysed in this experiment, namely *KfGI* and *KfZTL*. In 10 out of the 12 tissue samples analysed, *KfZTL* transcript levels were up-regulated. In the other two samples (LP3 dark and LP10 light), there was no significant difference from the wild type. *KfGI* was only detected as being up-regulated in *KfCDF2_FL* in three samples, namely LP1 light and LP10 light and dark. *KfGI* transcript levels were down-regulated in the meristem in the dark, and in the stem in the light and dark, but showed no difference relative to the wild type in the remaining 6 samples. This could suggest that *KfCDF2* is able to directly or indirectly induce *CO*, as opposed to this induction requiring a mechanism involving *GI*.

Table 7.3. Summary of flowering pathway gene transcript abundance changes relative to wild type in *KfCDF2_FL_14C*.

CDF2	LP1		LP3		LP6		LP10		Meristem		Stem	
FL 14C	L	D	L	D	L	D	L	D	L	D	L	D
CDF2												
GI												
CO25												
CO46												
FT12												
FT18												
AP1												
SEP												
ZTL												
SOC1												
FD												
FUL												

Green indicates increased expression, whilst red shows a decrease. Purple indicates levels were not significantly different from those in the wild type.

7.2.3.2 Changes in the transcript abundance of flowering pathway genes in *KfCDF2_RNAi_19A*

With *KfCDF2_RNAi_19A*, the majority of tissues at the majority of times did show a reduction in the *KfCDF2* transcript level, consistent with the 35S promoter driving the hairpin RNA RNAi transgene in this line. The exceptions were LP1 in the light and dark, and the meristem and stem at the end of the light, which all showed an increased transcript level for *KfCDF2* compared to the wild type despite the presence of the *KfCDF2_RNAi* transgene (Table 7.4).

In this RNAi knockdown line, many of the other flowering pathway genes were again up-regulated in all samples, including *KfCO25, KfCO46, KfAP1* and *KfSOC1*, which are all involved in promoting flowering. *KfZTL* was also over-expressed relative to the wild type in every tissue and time-point, except for at the end of the light period in LP10. On the other hand, *KfGI* only showed over-expression in the light period for each of the sampled leaf pairs. *KfGI* also showed down-regulation in the stem and during the dark in the meristem. In the dark in LP1, LP3 and LP6 there was no difference to wild type for *KfGI*, whilst in LP10 there was a decrease in levels in the dark.

In contrast to the finding in *KfCDF2_FL_14C,* where *KfFUL* was consistently over-expressed, in the *RNAi_19A* line it showed no difference to wild type in any sample. This may therefore suggest that *KfCDF2* has a direct or indirect role in inducing *KfFUL* expression, with excess *KfCDF2* leading to higher *KfFUL* levels.

As in *KfCDF2_FL*, *KfSEP* was also up regulated in the light, but down regulated in the dark, consistently in all sampled tissues/ organs. *KfFT12* was only found to be up-regulated in the stem and meristem relative to wild type, whereas in most other tissues it was found to be at the same level as in the wild type, or even down-regulated in

LP1 and LP3 in the dark (Table 7.4). However, it must be remembered that increased transcript levels in the stem and meristem are likely to be the most important areas in terms of flowering, as this is where the mobile FT is required to induce flowering and may even induce its own expression in the stem and meristem as the FT protein becomes more abundant. *KfFT18* on the other hand was generally under-expressed or expressed at the same level in the various tissues/ organs of the RNAi line. Over-expression of the KfFT18 transcripts was only detected in LP1 at the end of the dark and LP10 in the light. *KfFD* expression in the meristem was down-regulated in the RNAi line, but again this was only a small change relative to the WT (Appendix S7.11). LP1 in the light and dark, LP3 and the stem in the dark, and LP6 in the light all showed increased transcript levels for *KfFD*, with the largest increase being detected in the stem.

Table. 7.4. Summary of changes in the transcript abundance levels of flowering pathway genes relative to wild type in the *KfCDF2_RNAi_19A* transgenic line

CDF2	LP1		LP3		LP6		LP10		Meristem		Stem	
RNAi 19A	L	D	L	D	L	D	L	D	L	D	L	D
CDF2												
GI												
CO25												
CO46												
FT12												
FT18												
AP1												
SEP												
ZTL												
SOC1												
FD												
FUL												

Green indicates increased expression, whilst red shows a decrease. Purple indicates levels do no differ from wild type.

As part of follow-up work aimed at determining the downstream target genes that *KfCDF2* binds to the promoters of, transgenic lines that over-expressed a HA-tagged

version of *KfCDF2_FL* were produced in *K. laxiflora* in the final 12 months of this project. *K. fedtschenkoi* and *K. laxiflora* are very closely related and very similar phenotypically, but *K. laxiflora* has been found to set viable seed when self-pollinated and so is believed to be a more amenable model species for further transgenic experiments. The two species have very similar genomes, although *K. laxiflora* is a tetraploid with a ~490 Mb genome whereas *K. fedtschenkoi* is a diploid with a 246 Mb genome.

Interstingly, the impact of overexpressing *HA-KfCDF2* in *K. laxiflora* appeared to be much stronger than when the same gene was over-expressed in *K. fedstchenkoi*. All of the *K. laxiflora* lines regenerated with the *HA-KfCDF2_FL* over-expression construct flowered early in long days in the greenhouse, indicating that the early flowering in long days phenotype displayed 100% penetrance for all screened and confirmed HA-*KFCDF2* overexpression lines in *K. laxiflora*. Samples have been collected from these lines, and hopefully the next step will be Chromatin Immunoprecipitation followed by high-throughput sequencing on the Illumina system (ChIP-Seq). This work should hopefully reveal exactly where this TF binds in the genome, and thus which downstream target genes are regulated by *KfCDF2*.

7.3 Discussion

As well as playing a role in the circadian optimisation of CAM (see Chapters 3 and 6), *KfCDF2* has also been identified here as playing a role in the photoperiodic regulation of the short-day-dependent flowering response in *K. fedtschenkoi* and *K. laxiflora*. *KfCDF2* over-expresser transgenic lines were found to flower constitutively in 16:8 LD greenhouse conditions (long days); conditions which have never been found to mediate flowering in wild type *K. fedtschenkoi*. The over-expresser lines were also shown to flower more rapidly than the wild type under both 12:12 and 8:16 light/dark conditions. They produced more flowers, and, visually at least, appeared to have fewer leaves, appearing to invest less energy into vegetative growth, and more into the reproductive effort.

Interestingly though, it appears that the quality and/or quantity of light affects the ability of the *KfCDF2* transgenic lines to flower. It was found that low levels of blue, red and far-red light in the visible spectrum and low light intensity, leading to flowering inhibition in the transgenic lines. This light quality and quantity effect did not affect the wild type though, which displayed a normal flowering response in 8:16 LD (short days). DOF transcription factors have been shown to play roles in phytochrome signalling, involving red/far-red light. An example of this is *COG1*, which when overexpressed leads to hyposensitivity to red and far-red light (Park *et al.,* 2003). This opens up the possibility that *KfCDF2* also plays a role in the phytochrome A and B signalling pathways. Phytochrome A mutants have also been shown to have delayed flowering in low irradiance (Bagnall *et al.,* 2001). Furthermore, because there was also less blue light in the growth room light spectrum, this may have led to ZTL and its subsequent complexes becoming less stable, which could lead to flowering inhibitors not being degraded. In relation to light quantity

during the day, the intensity was also found to be much lower in the growth room where the *KfCDF2* transgenic lines did not flower, compared to the green house and growth chambers where flowering was successful.

It is important to note in this context that SDPs have been found in general to be less sensitive to light quality and quantity (Park *et al.,* 2003). Instead, they are often characterised as being more sensitive to the length of the dark period, although this is not known specifically for the short day response in *K. fedtschenkoi*. Low irradiance levels will of course affect photosynthesis, and these plants have already been shown in the Chapter 6 to achieve reduced levels of dark CO_2 fixation relative to the wild type. This could therefore mean that they have fewer resources available to fuel the development of the floral meristem, suggesting that metabolic changes as a result of perturbations in photosynthesis could be the root cause of the inhibition of flowering in the *KfCDF2* transgenic lines under the growth room conditions used in the study.

There is a wealth of knowledge regarding the genetic and signal-transduction pathways involved in the regulation of flowering time in Arabidopsis, but much less is known about the equivalent pathways in short-day plants. Samples of all above-ground parts of the plants were therefore collected throughout the flower-induction process under short day conditions, and the transcript levels of flowering pathway genes were investigated in LP1, LP3, LP6, LP10, stem, and meristem, and later on buds and flowers. From this, gene transcript abundance could be monitored in the weeks leading up to the opening of the first flowers.

7.3.1 Flowering progression in wild type *K. fedtschenkoi*

The use of RT-qPCR to study the regulation of the steady-state transcript abundance of *KfCDF2* in wild type *K. fedtschenkoi* found that there was high transcript abundance for *KfCDF2* in long days in LP6 and LP10, but low transcript levels in other parts of the plant. This suggested that the increase in transcript levels in LP6 and LP10 may be due to CAM induction, where the *KfCDF2 TF* plays a role in the circadian control of CAM (Chapter 6). Once wild type plants entered short day conditions that induce flowering, the transcript abundance of *KfCDF2* was elevated in other tissues. From week 5, *KfCDF2* showed a large increase in the transcript abundance of *KfCDF2* in the stem and meristem, suggesting KfCDF2 may have different regulatory roles in different parts of the plant. In Arabidopsis, it has been demonstrated that *CDFs* repress *CO* expression in the leaves (Fornara *et al.,* 2009), but in the meristem *CDFs* may have a different role, possibly depending on which protein it is interacting with.

It is possible that *KfCDF2* functions as a promoter of flowering rather than a repressor in *K. fedtschenkoi*, as a sequence alignment of the amino acid sequences of the *KfCDF2* genes and the closest Arabidopsis ortholog revealed that they share only 38.8 % sequence identity. In addition, once the wild type plants had been transferred to short day conditions, the transcript abundance of *KfCDF2* dropped in older leaves, which could possibly be due to levels being reduced by GI and ZTL. This is supported by transcript data for both genes, showing increased transcript levels after a week of short day inductive conditions, which correlated with the reduction in *KfCDF2* transcript levels in the older leaf pairs. It would therefore be interesting to investigate the impact on CAM in *K. fedtschenkoi* that are progressing to flower in

short-days. If *KfCDF2* is down regulated, are other CAM genes also down regulated? Could flowering and reproduction be being prioritised over CAM?

The two *CO* sequences in *K. fedtschenkoi* shared 84.5% pairwise identity in an alignment of their nucleotide sequences, but showed peak transcript levels at different times of day. *KfCO25* tended to peak at the end of the light period, whereas *KfCO46* peaked during the dark. It therefore seems feasible that the ratio of expression between these two genes that leads to flowering could be the important factor for the induction of flowering in *K. fedtschenkoi*. *KfCO46* was also detected at high levels in buds, suggesting *KfCO46* could also function in the regulation of floral development, whereas *KfCO25* may perform more typical CO functions including mediating the induction of *FT* in short days.

For the two *FT* genes identified, the regulation of the transcript level of *KfFT18* was consistent with it acting as the antiflorigen, whilst *KfFT12* regulation was consistent with it functioning as the *K. fedtschenkoi* florigen; despite the homology to the Arabidopsis FT genes suggesting the opposite functions. *KfFT12* was found to increase in tissues the longer they were in short day 8:16 LD conditions, and was also found in developing buds, where *KfFT18* was barely expressed. The two genes also peaked at different times, which could suggest they regulate each other. *KfFT18* peaked during the light, suggesting that in long days lots of inhibitor is expressed preventing flowering. By contrast, *KfFT12* peaked during the dark, and so in longer nights more of the florigen would be able to be synthesised, leading to induction of flowering and could repress *KfFT18*.

KfAP1 was transcribed in all tissues even in non-flowering conditions, which is also the case in Arabidopsis (TAIR; www.arabidopsis.org), suggesting it does play other roles in the plant on a day-to-day basis beyond its best characterised role in the

regulation of floral development. However, *KfAP1* transcript levels increased by week 5/6 in all tissues. This is a developmental point at which the floral meristem had begun to develop, and, based on the known functions of its ortholog in Arabidopsis, *KfAP1* transcripts would be a key element in the development of the floral meristem (Teper-Bamnolker *et al.,* 2005). *KfAP1* also increased in transcript abundance in buds, suggesting this gene plays the same role in the regulation of flowering in *K. fedtschenkoi*, as it does for Arabidopsis, which is controlling sepal and petal differentiation and development (Mandel *et al.,* 1992; Kaufmann *et al.,* 2010).

KfSEP is also an important gene for flowering, and was repressed during the light period in non-inductive conditions. Under short-day inductive conditions, light period transcript levels for *KfSEP* were detected. *SEP2* and *SEP3* have been shown to function redundantly with *SEP1* (Pelaz et al., 2000), and are crucial for the transition to flower development in the meristem in Arabidopsis (Teper-Bamnolker *et al.,* 2005). Levels of *KfSEP* in *Kalanchoë* were found to be high in the meristem and also in buds, again suggesting this gene functions in a similar way to its Arabidopsis ortholog.

KfZTL appeared to play a role in the final floral development signal, as in all tissues at week 7, right before visible flowers were produced, there was a large increase in the transcript level of *KfZTL*. A high level of *KfZTL* transcript was detected in buds that were starting to open.

Despite *KfSOC1* being a flowering gene, it has also been shown recently to play a role in stomatal control (Kimura *et al.,* 2015). This could explain why in leaf pairs 6 and 10, levels were high under non-inductive long-day conditions. Once plants were transferred to short-days and the progression to flowering had commenced, the levels of this transcript dropped. This could again suggest that flowering may be prioritised

over CAM. It would be of interest to examine stomatal behaviour in leaves during the short-day induction of flowering in *K. fedtschenkoi*.

KfFD transcript regulation did not correlate well with FD's proposed function in Arabidopsis. FD has been shown to interact with FT to induce flowering. *KfFT12* was not really detected until at least week 5 in all tissues, whereas FD expression greatly drops after week 5. It may indicate that only small levels of the two proteins are needed to induce flowering, or could suggest a different pathway is used. Another suggestion could also be that as *KfFT18* has the highest homology to the *Arabidopsis* florigen, then *KfFD* may in fact interact with that instead, and so may have taken on a flowering repression role.

KfFUL transcript regulation displayed a similar pattern of transcript levels to its Arabidopsis counterpart. It was barely detected in non-inductive long day conditions, but increased in all tissues under short days from week 5, and also increased as the buds developed, supporting its proposed role as a gene that promotes flowering. Interestingly, in *Arabidopsis*, *LFY* is a very important gene for flowering, as it was shown to be both necessary and sufficient to induce both flowers and leaves (Blazquez *et al.,* 1997). A *KfLFY* gene candidate (KF07545) was found in the *K. fedtschenkoi* genome, and was amplified successfully from genomic DNA, but it was not detected as an expressed transcript in both an existing database of RNA-seq quantitative transcriptomic data, and also after 40 cycles of RT-qPCR. This could suggest that an alternative gene performs the role of *LFY* in *K. fedtschenkoi*, or it may be that it can achieve its required function in *K. fedtschenkoi* from a very low transcript level that could not be detected with the methods used. Unfortunately, the role of *KfLFY* in the *K. fedtschenkoi* flowering pathway could not be determined, and

so this may suggest that *KfAP1* is able to function alone to initiate the development of flowers in this species.

7.3.2 The impact of changes in *KfCDF2* expression on the regulation of genes associated with the flowering pathway

With the regulation of flowering pathway genes in wild type in response to short days established as a baseline, it was then of interest to investigate the same genes for possible differential regulation in the *KfCDF2_FL* line that was able to initiate flowering in good quality greenhouse light under long days, without receiving a short day flowering signal. As *KfCDF2_RNAi_19A* also showed a more rapid stem hooking phenotype in 12:12 LD conditions relative to wild type, it seemed that this down-regulated *KfCDF2* transgenic lines was also able to flower more rapidly in response to short days, but only if given the initial flowering stimulus, which the *KfCDF2_FL* lines no longer require.

In both the *KfCDF2_FL* and *RNAi* line, many flowering pathway genes were up-regulated constantly; in particular *KfCO25, KfCO46* and *KfAP1*. This data provides a molecular-genetic basis to explain why the *KfCDF2* over-expresser line could flower more rapidly, and also suggests why, under 12:12 conditions, the *KfCDF2_RNAi* line was also primed to flower. For *KfCDF2_FL_14C*, there were many more genes up regulated in many more tissues than in the *RNAi_19A* line. Constitutive overexpression of *KfFUL* was also detected in the FL line, and constantly high transcript levels of *KfSEP* were detected in the light. The accumulation of *KfFUL*, *KfSEP3* and *KfAP1* in the meristem is the key step in induction of the floral meristem in *K. fedtschenkoi*. *KfCDF2_FL* had a constantly high level of these transcripts, which suggests that the *KfCDF2_FL* over-expresser line is constantly receiving a flower-

inducing signal, and that plants are simply waiting until they have accumulated sufficient storage carbohydrate/ lipid and/ or protein reserves to be able to initiate flowering.

By contrast, in the *KfCDF2_RNAi_19A*, transcript abundance for *KfFUL* was similar to the wild type, which is consistent with the fact that these lines did not initiate flowering in 16:8 LD conditions (long days). In the *KfCDF2_RNAi* line, *KfCDF2* expression was found to increase in the meristem and stem in the light, but this was not sufficient to signal an increase in *KfFUL*, even though there was an increase in *KfFT12*. This suggests that a threshold level of *KfCDF2* is required in order to induce *KfFUL* expression, which would in turn initiate the transition to flowering. This could also help to explain why some of the independent *KfCDF2_FL* over-expresser lines only progressed as far as producing hooked stems under 16:8 greenhouse conditions, but did not actually succeed in flowering. This could be due to these lines having a level of *KfCDF2* over-expression that was not quite sufficient to cause an increase in *KfFUL* expression to the required level.

In the *KfCDF2_RNAi* lines, *KfSOC1* was over-expressed constantly in all tissues sampled. This could be a further reason why these lines are not flowering independently of the photoperiod, because in non-inductive long day conditions, *KfSOC1* was over-expressed, which may have prevented the progression to inflorescence meristem initiation (Lee *et al.,* 2010).

Finally, it was surprising that both *KfCDF2* transgenic lines lacked a detectable level of transcripts for the *KfLFY* gene that was identified in the genome. This suggests that *KfLFY* was not expressed in *K. fedtschenkoi* under the conditions studied here using the RT-qPCR technique. It would appear from these findings that *KfLFY* does not play a role in the regulation of floral development in response to short days in *K.*

fedtschenkoi. This is certainly an area that merits further work. A more comprehensive RT-qPCR or quantitative RNA-seq study should be undertaken to determine whether or not the *KfLFY* gene is expressed under any conditions and / or in any specific tissue / organ during the development of a *K. fedtschenkoi* plant. It is noteworthy in this context that an assembled *K. laxiflora LEAFY* transcript was detected in a PASA (Program to Assemble Spliced Alignments) assembly of *K. laxiflora* transcripts that were assembled using Illumina Hi-Seq reads from a wide range of tissues, organs and developmental stages (Won Cheol Lim, Hengfu Yin, Xiaohan Yang, John Cushman and James Hartwell, unpublished findings). This suggests that *KlLFY* was being expressed in some of the wide range of *K. laxiflora* tissues and developmental time points that were sampled for the Illumina sequencing that led to the *K. laxiflora* PASA transcriptome assembly. These findings further emphasise the importance of future work to identify which tissues/ organs and/ or developmental stages possess detectable levels of *KlLFY* or *KfLFY* expression in *K. laxiflora* and *K. fedtschenkoi* respectively.

7.4 Conclusions

There were many similarities between the patterns of gene regulation observed during the transition to flowering in short days for *K. fedtschenkoi* studied here and previous studies with Arabidopsis. This is interesting considering that Arabidopsis is a LDP and *K. fedtschenkoi* is a SDP. The transcript abundance profiles measured here indicate that many flowering pathway genes are likely to carry out similar functions in both species, including *KfAP1, KfCO* and *KfFUL*. Other flowering pathway genes appeared to function in a different way in *K. fedtschenkoi* relative to the known function and pattern of regulation in Arabidopsis, including *KfFD* and *KfLFY*; with the latter gene not being detected as expressed in any of the sampled tissues or time-points. *KfCDF2* was one of the genes that was found to conduct a different function, as in LDPs *CDF2* overexpression has been shown to delay flowering time, rather than initiating and accelerating it (Fornara *et al.,* 2009).

Much more work needs to be conducted in this area to establish the exact functions of many of these flowering genes, and also establish changes associated with day length and also CAM, which *KfCDF2* has been demonstrated to play a role in regulating (Chapter 6). Furthermore, the findings reported in this chapter, and the recently published findings of Kimura *et al.* (2015), suggest that *KfSOC1* could be involved in stomatal regulation in CAM performing leaves of *K. fedtschenkoi* where it was up-regulated, and this TF also merits detailed further investigation through the development of transgenic over-expresser and RNAi lines, and ChIP-Seq analysis. The *KfCDF2* transgenic lines were found to be more sensitive to light quality or quantity than wild type lines, as they did not flowering in growth room short day conditions under which the wild type did successfully flower. It is interesting in this context that some *DOF* factors have already been shown to be involved in light

sensing and signalling pathways in Arabidopsis (Park *et al.,* 2003). Further work needs to be carried out, with careful control of both the quantity and quality of the light, in order to determine whether light quality or quantity is important for the *KfCDF2* transgenic lines to flower early. If light quality is most important then experiments using different LED lamps that emit very specific wavelengths of light should permit a precise dissection of the optimal wavelength(s) of light required for the induction of flowering through a *KfCDF2* signalling pathway in response to short days.

ChIP-Seq analysis of the genome wide promoter and other gene regulatory regions bound by *KfCDF2* will also be a critical step towards understanding which genes are bound to and regulated by *KfCDF2*. These results will shed more light on the direct functions of *KfCDF2*. It will also be interesting to investigate the upstream promoter region of the *KfCDF2* gene itself, in order to identify the important promoter motifs required for its induction in leaf pair 6 relative to leaf pair 1, and its light/ dark and circadian clock control. This could be pursued using promoter deletion analysis with a series of *KfCDF2* promoter::luciferase reporter constructs. Luciferase imaging of stable transgenic lines carrying different length fragments of the *KfCDF2* promoter should allow the minimal promoter for leaf pair 6 induction, and light/ dark and circadian control to be identified.

It was also interesting to discover during the last few months of the practical work on this project that the newly generated *HA-KfCDF2* over-expresser lines in *K. laxiflora* all progressed to flowering in 16:8 long day conditions in the greenhouse. This suggested that *K. laxiflora* may be more susceptible to over-expression of *KfCDF2* than *K. fedtschenkoi*, as only the strongest *KfCDF2* over-expresser lines in *K. fedtschenkoi* flowered constitutively in long day greenhouse conditions. However, it

is possible that all of the more recently generated *K. laxiflora* lines over-express *KfCDF2* more strongly that most of the previous population of *K. fedtschenkoi KfCDF2* over-expresser lines and further work must now be undertaken to characterise the level of transgene expression in the new *HA-KfCDF2* over-expresser lines of *K. laxiflora*. However, it remains possible that flowering in *K. laxiflora* was more susceptible to elevated levels of *KfCDF2*. It is thought provoking to speculate that this may be in some way be linked to the fact that *K. laxiflora* is able to develop viable seed following self pollination, whereas *K. fedtschenkoi* fails to form viable seed due to a mutation in the *LEAFY-COTYLEDON1* (*LEC1*) gene (Garces *et al.,* 2007). The *LEC1* mutation found in several plantlet-forming species of *Kalanchoë* causes a C-terminal truncation of the *LEC1* protein, and results in a failure of developing seed to achieve dehydration tolerance in the final stages of seed maturation. Whether or not this mutation also impacts on the flowering pathway and the role of *KfCDF2* in that signalling pathway is an interesting area for future investigation, which is readily tractable via a direct comparison of flowering pathway signalling in *K. fedtschenkoi*/ *K. laxiflora,* as the two species possess a mutated *LEC1* and a functional *LEC1,* respectively.

Chapter 8 - General Discussion and Future Work

8.1 General Discussion

The main aim of this PhD was to characterise the *in planta* function of transcription factors (TFs) that had been identified from quantitative RNA-seq analysis using the Applied Biosystems SOLiD sequencer as having a potential role in the circadian optimisation of CAM. These genes were chosen by virtue of the fact that their closest orthologs in the genome of Arabidopsis were well-characterised DNA-binding transcription factors. They also underwent diurnal oscillations of their steady-state transcript abundance in LD cycles in CAM leaves of *K. fedtschenkoi*, and continued to oscillate with a robust circadian rhythm under free-running LL conditions. This method of identifying genes with a potential role in the circadian clock control of the CAM pathway has proven successful. All three genes (*KfMYB439, KfCIB1* and *KfCDF2)* were characterised with respect to their *in planta* function using both RNAi knock-down and FL over-expresser transgenic lines of *K. fedtschenkoi*. In all cases, the perturbation of the transcript abundance of these TFs led to perturbations in CAM and its daily and circadian clock control. Manipulating the level of all three genes caused widespread but small changes in the steady-state transcript levels of a range of CAM-associated genes.

Based on the findings presented in this thesis, *KfMYB439* most likely functions close to the core circadian clock, which fits with this gene having been identified as a member of the *RVE* family, which includes the core clock genes *CCA1* and *LHY (*Farinas *et al., 2011)*. In addition to the *KfMYB439* transgenic lines displaying widespread changes in the transcript levels of CAM-associated genes, they also showed a change in the timing of their daily patterns of CO_2 fixation. Changes in expression of *KfMYB439*, both overexpression and RNAi knockdown, led to less

Chapter 9 - Referecnes

ABE, M., KOBAYASHI, Y., YAMAMOTO, S., DAIMON, Y., YAMAGUCHI, A., IKEDA, Y., ICHINOKI, H., NOTAGUCHI, M., GOTO, K. & ARAKI, T. 2005. FD, a bZIP protein mediating signals from the floral pathway integrator FT at the shoot apex. *Science,* 309**,** 1052-6.

ACEVEDO, E., BADILLA, I. & NOBEL, P. S. 1983. Water Relations, Diurnal Acidity Changes, and Productivity of a Cultivated Cactus, Opuntia ficus-indica. *Plant* Physiol*,* 72**,** 775-80.

ADAMS, P., NELSON, D. E., YAMADA, S., CHMARA, W., JENSEN, R. G., BOHNERT, H. J. & GRIFFITHS, H. 1998. Growth and development of Mesembryanthemum crystallinum (Aizoaceae). *New Phytologist,* 138**,** 171-190.

AHN, N. G., SHABB, J. B., OLD, W. M. & RESING, K. A. 2007. Achieving in-depth proteomics profiling by mass spectrometry. *ACS Chem Biol,* 2**,** 39-52.

ALBANI, D., HAMMONDKOSACK, M. C. U., SMITH, C., CONLAN, S., COLOT, V., HOLDSWORTH, M. & BEVAN, M. W. 1997. The wheat transcriptional activator SPA: A seed-specific bZIP protein that recognizes the GCN4-like motif in the bifactorial endosperm box of prolamin genes. *Plant Cell,* 9**,** 171-184.

ANDERSON, C. M. & WILKINS, M. B. 1989a. Control of the circadian rhythm of carbon dioxide assimilation in Bryophyllum leaves by exposure to darkness and high carbon dioxide concentrations. *Planta,* 177**,** 401-8.

ANDERSON, C. M. & WILKINS, M. B. 1989b. Period and phase control by temperature in the circadian rhythm of carbon dioxide fixation in illuminated leaves of Bryophyllum fedtschenkoi. *Planta,* 177**,** 456-69.

ANDRONIS, C., BARAK, S., KNOWLES, S. M., SUGANO, S. & TOBIN, E. M. 2008. The clock protein CCA1 and the bZIP transcription factor HY5 physically interact to regulate gene expression in Arabidopsis. *Molecular Plant,* 1**,** 58-67.

BAENA-GONZALEZ, E., ROLLAND, F., THEVELEIN, J. M. & SHEEN, J. 2007. A central integrator of transcription networks in plant stress and energy signalling. *Nature,* 448**,** 938-U10.

BAGNALL, D. J. & KING, R. W. 2001. Phytochrome, photosynthesis and flowering of Arabidopsis thaliana: photophysiological studies using mutants and transgenic lines. *Australian Journal of Plant Physiology,* 28**,** 401-408.

BARTELS, D. & SUNKAR, R. 2005. Drought and salt tolerance in plants. *Crit. Rev. Plant Sci.,* 24**,** 23-58.

BAUMANN, K., DE PAOLIS, A., COSTANTINO, P. & GUALBERTI, G. 1999. The DNA binding site of the Dof protein NtBBF1 is essential for tissue-specific and auxin-regulated expression of the rolB oncogene in plants. *Plant Cell,* 11**,** 323-333.

BLASING, O., GIBON, Y., GUNTHER, M., HOHNE, M., MORCUENDE, R., OSUNA, D., THIMM, O., USADEL, B., SCHEIBLE, W. & STITT, M. 2005. Sugars and circadian regulation make major contributions to the global regulation of diurnal gene expression in Arabidopsis. *Plant Cell,* 17**,** 3257 - 3281.

BLAZQUEZ, M. A., SOOWAL, L. N., LEE, I. & WEIGEL, D. 1997. LEAFY expression and flower initiation in Arabidopsis. *Development,* 124**,** 3835-44.

BOHN, A., GEIST, A., RASCHER, U. & LUTTGE, U. 2001. Responses to different external light rhythms by the circadian rhythm of Crassulacean acid metabolism in Kalanchoe daigremontiana. *Plant Cell Environ.,* 24**,** 811-20.

Appendix.

Chapter 2 –

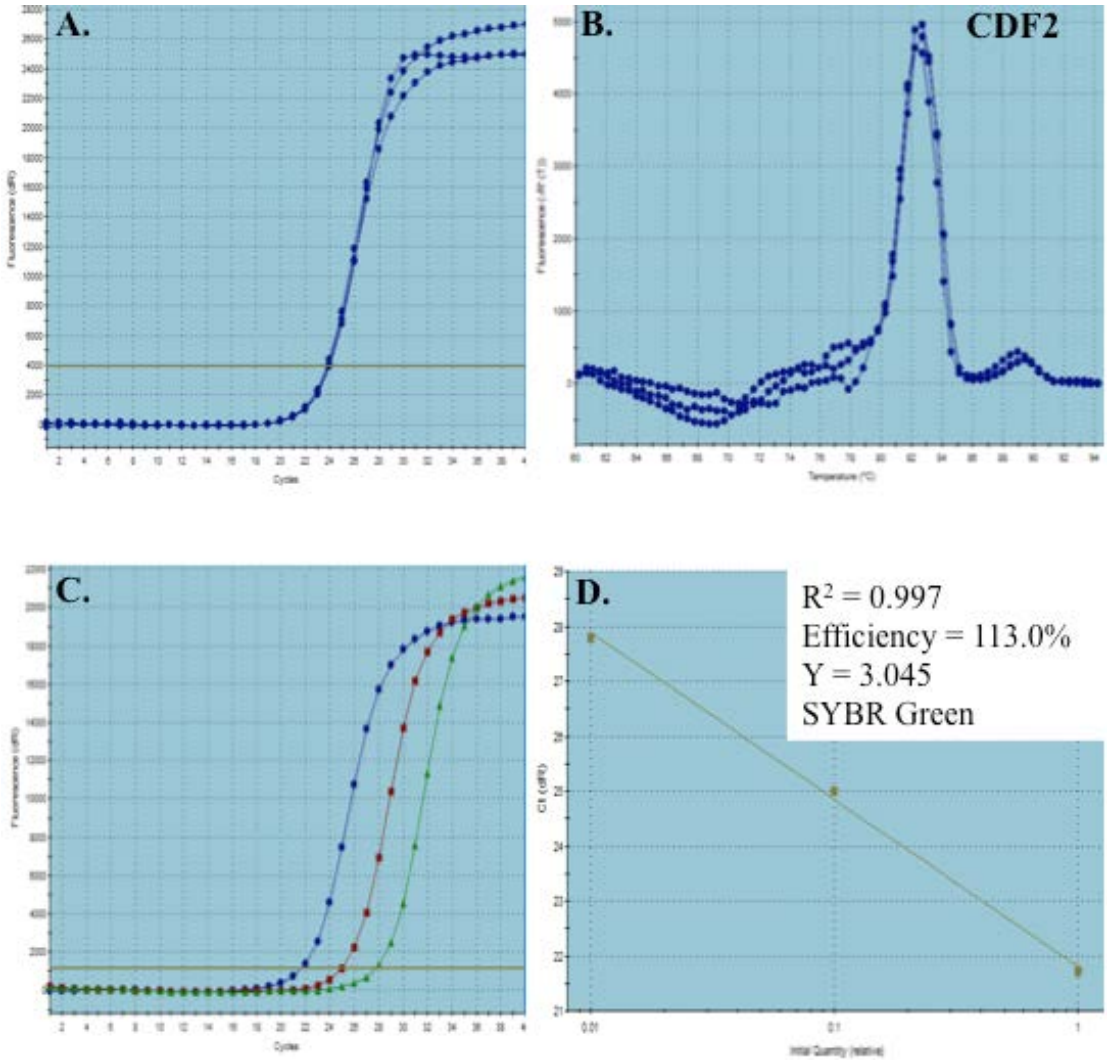

Figure S2.1. RT-qPCR graphs for primers of KfCDF2 used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red) and 100x dilution (green). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 113%.

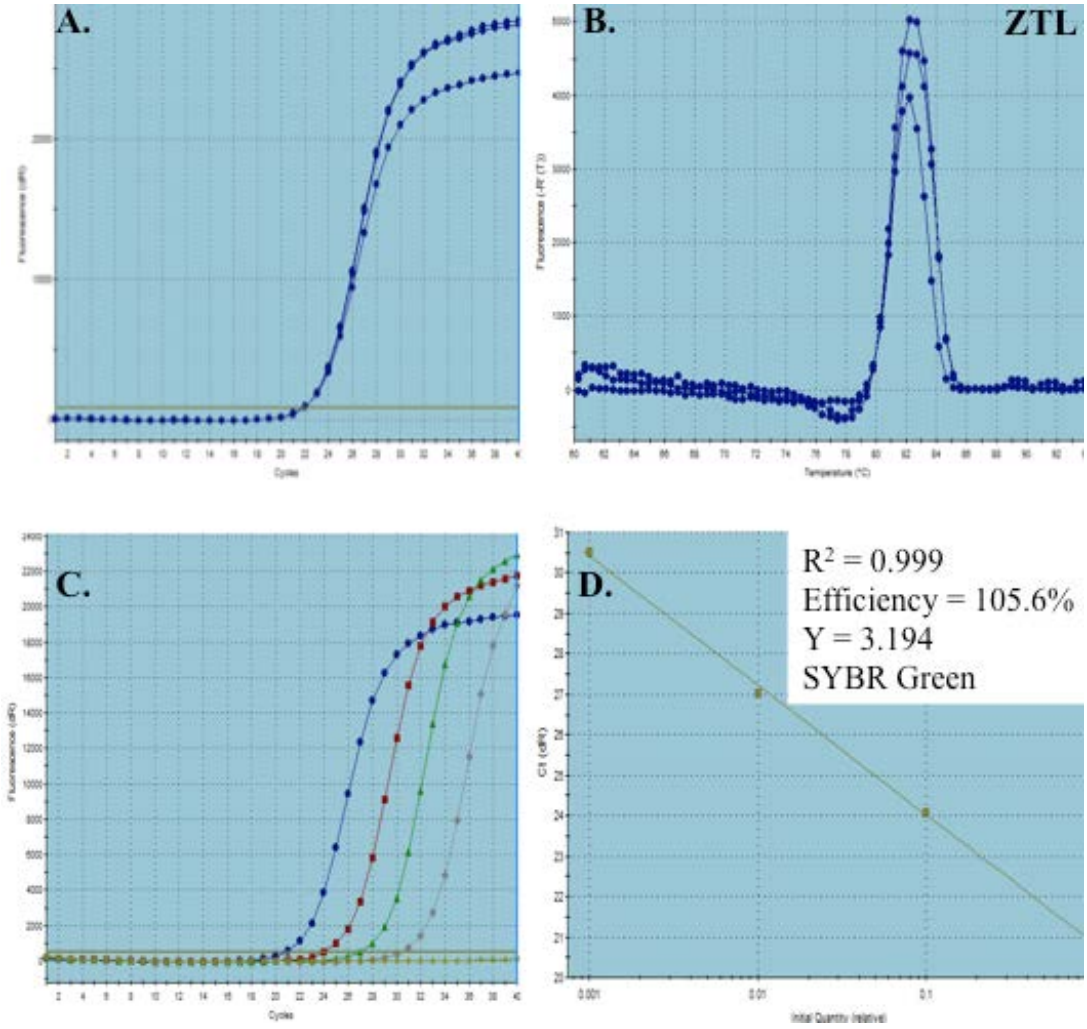

Figure S2.2. RT-qPCR graphs for primers of KfZTL used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red), 100x dilution (green) and 1000x dilution (grey). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 105.6%.

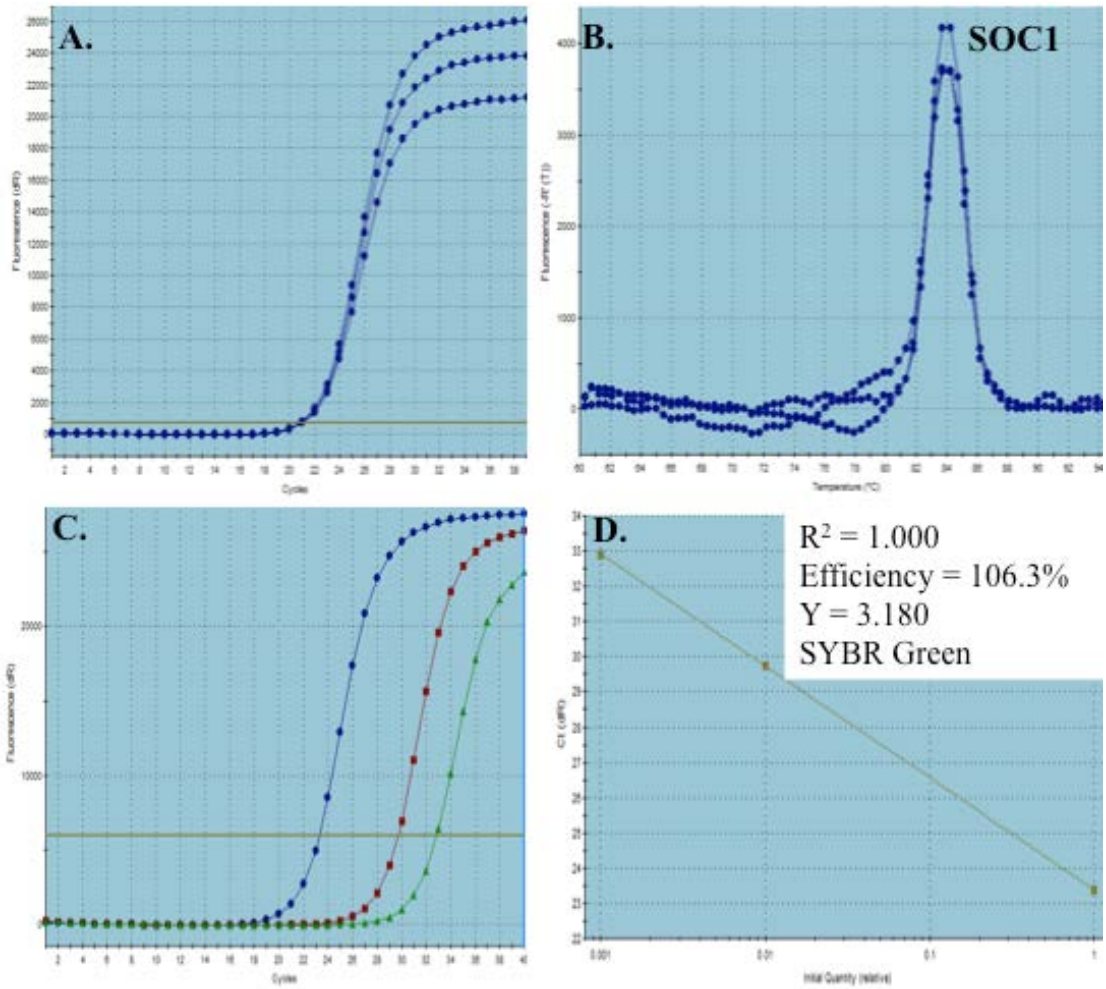

Figure S2.3. RT-qPCR graphs for primers of KfSOC1 used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red) and 100x dilution (green). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 106.3%.

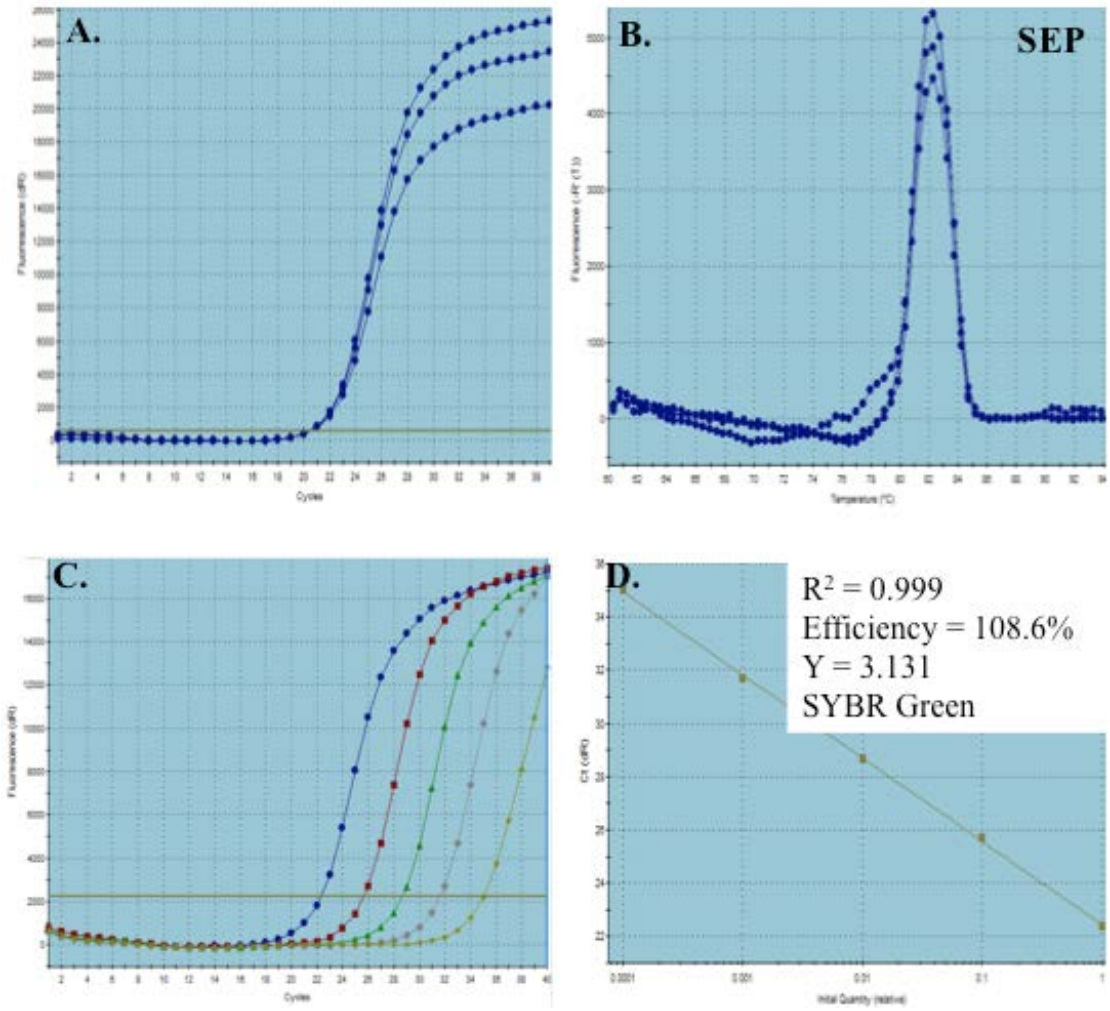

Figure S2.4. RT-qPCR graphs for primers of KfSEP used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red), 100x dilution (green), 1000x dilution (grey) and 10000x dilution (yellow). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 108.6%.

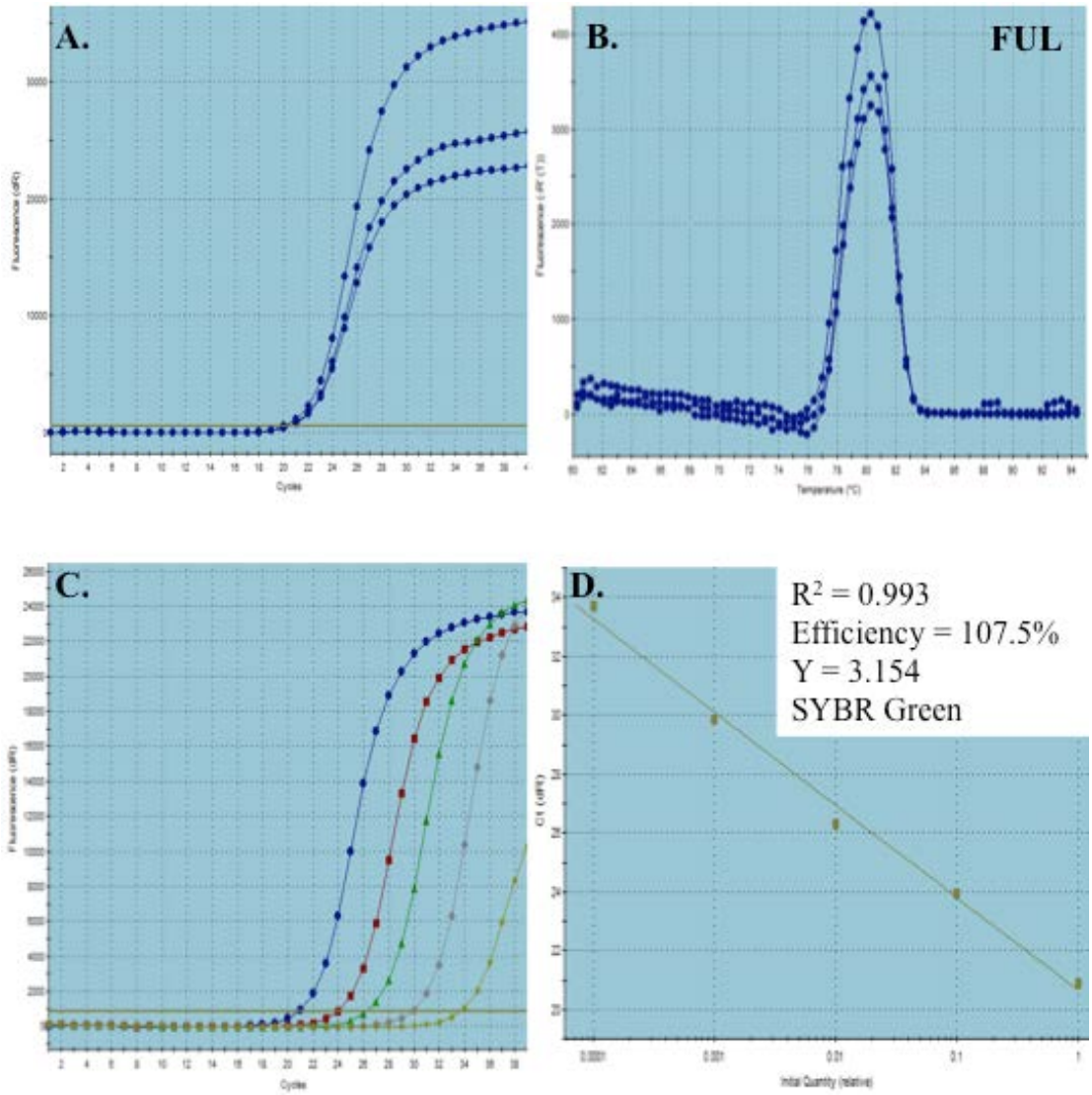

Figure S2.5. RT-qPCR graphs for primers of KfFUL used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red), 100x dilution (green), 1000x dilution (grey) and 10000x dilution (yellow). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 107.5%.

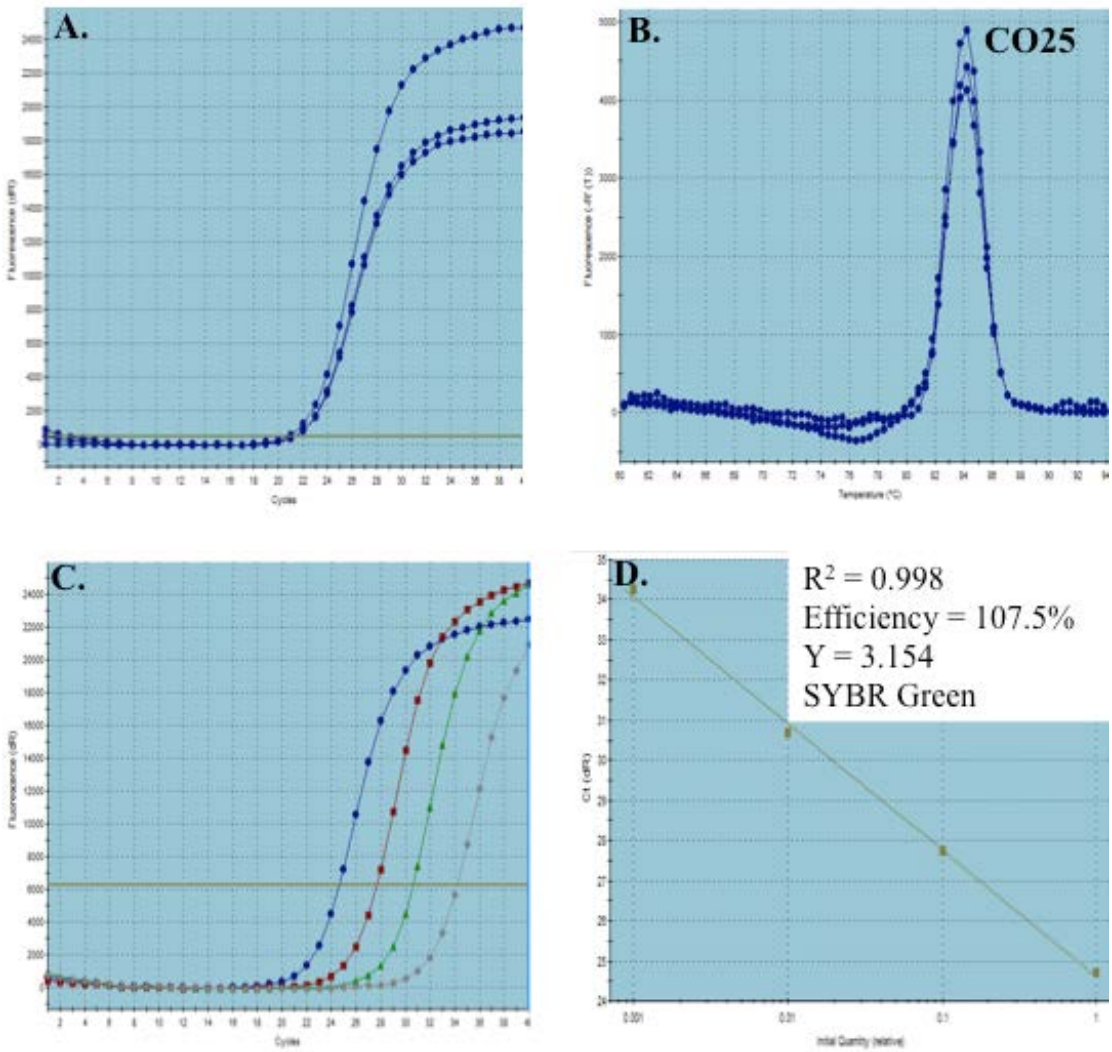

Figure S2.6. RT-qPCR graphs for primers of KfCO25 used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red), 100x dilution (green) and 1000x (grey). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 107.5%.varying dilutions

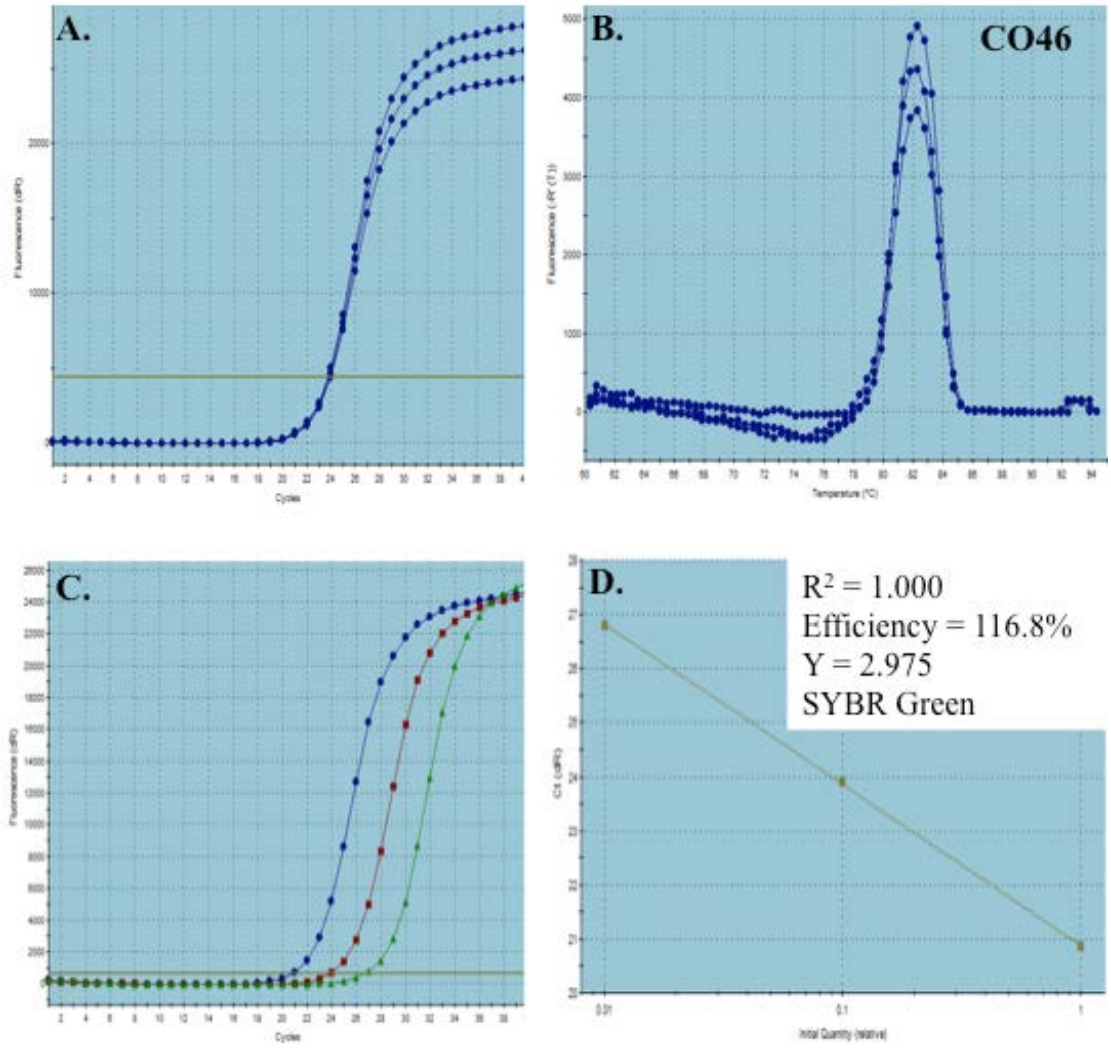

Figure S2.7. RT-qPCR graphs for primers of KfCO46 used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red) and 100x dilution (green). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 116.8%.

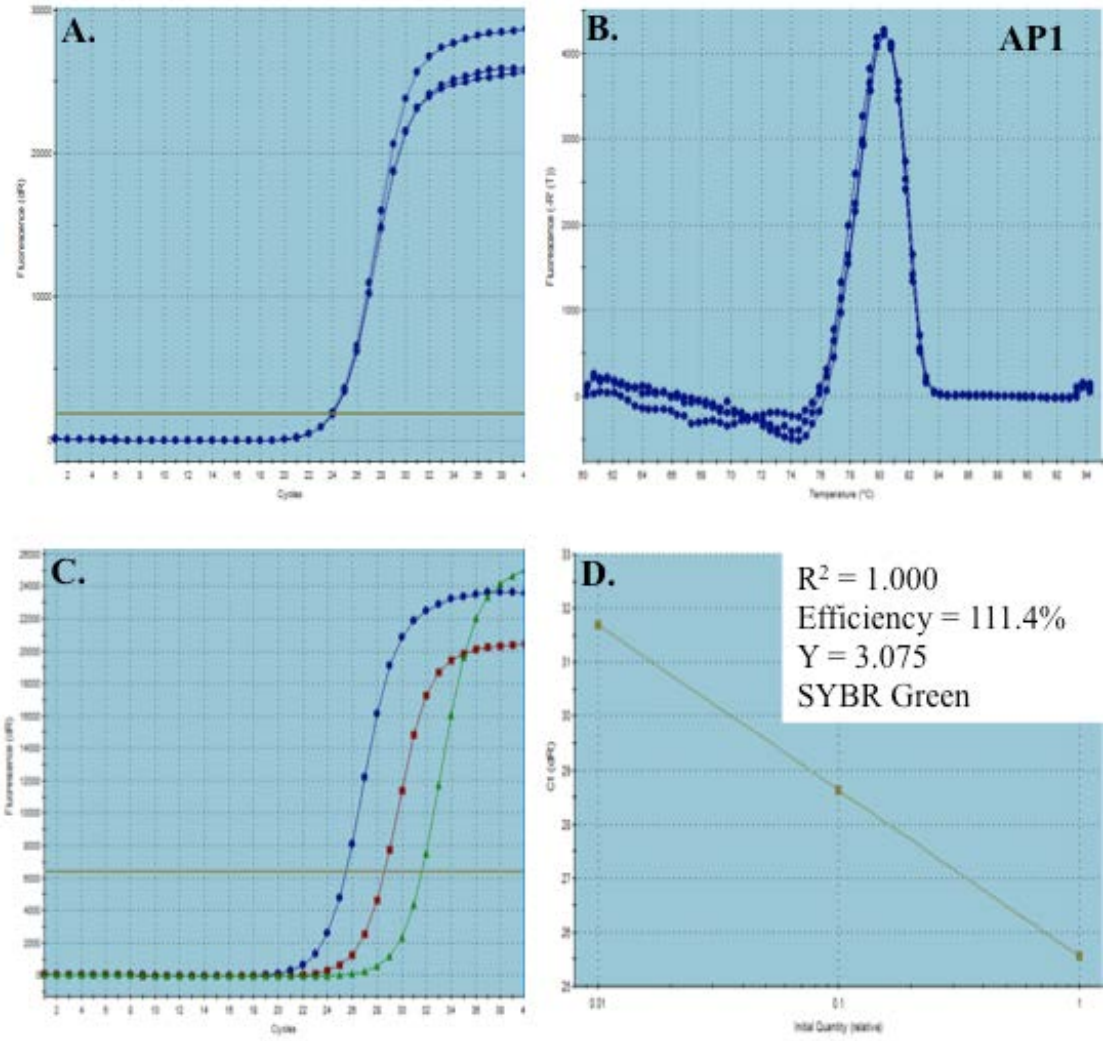

Figure S2.8. RT-qPCR graphs for primers of KfAP1 used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red) and 100x dilution (green). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 111.4%.

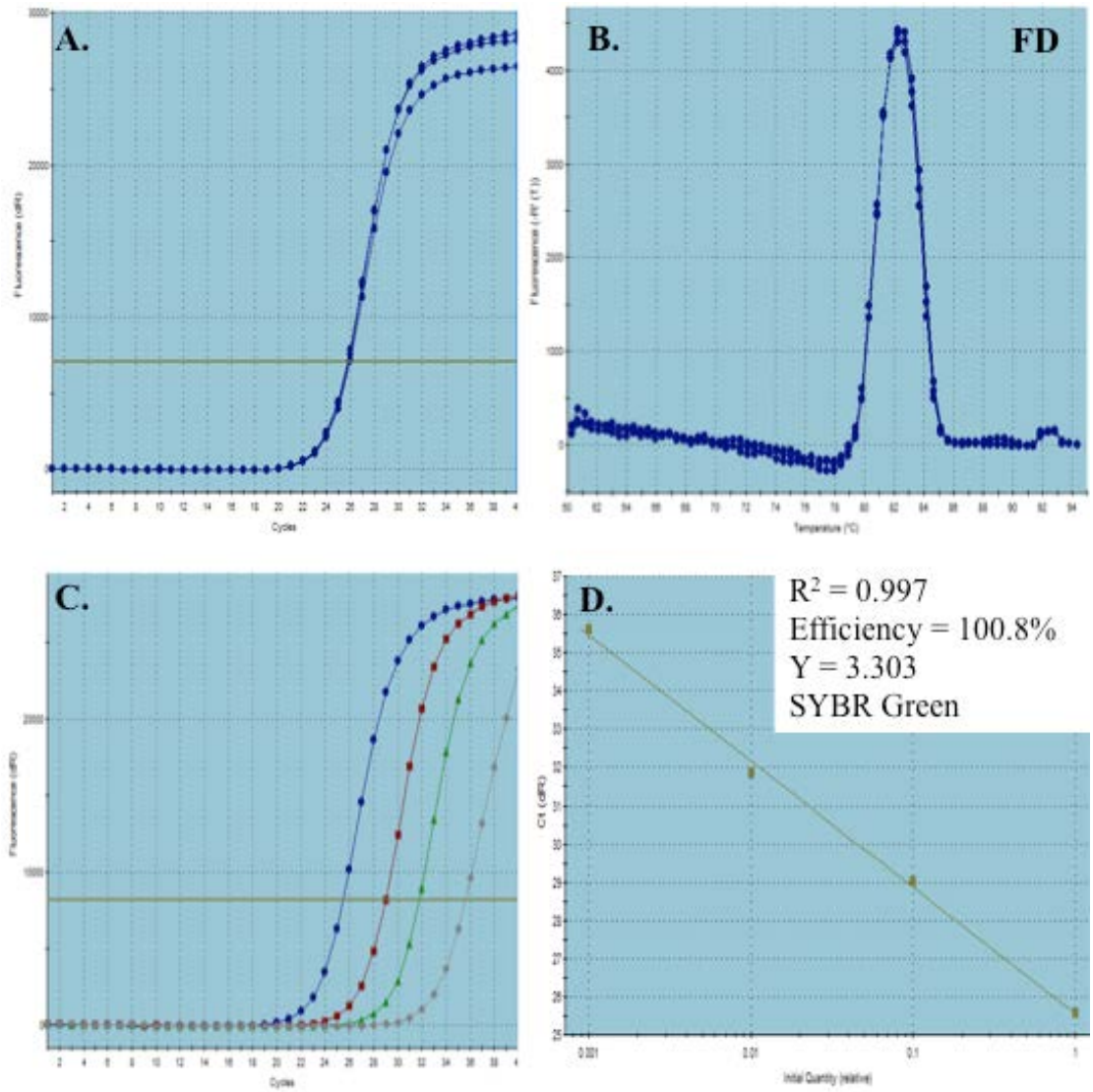

Figure S2.9. RT-qPCR graphs for primers of KfFD used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red), 100x dilution (green) and 1000x (grey). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 100.8%.

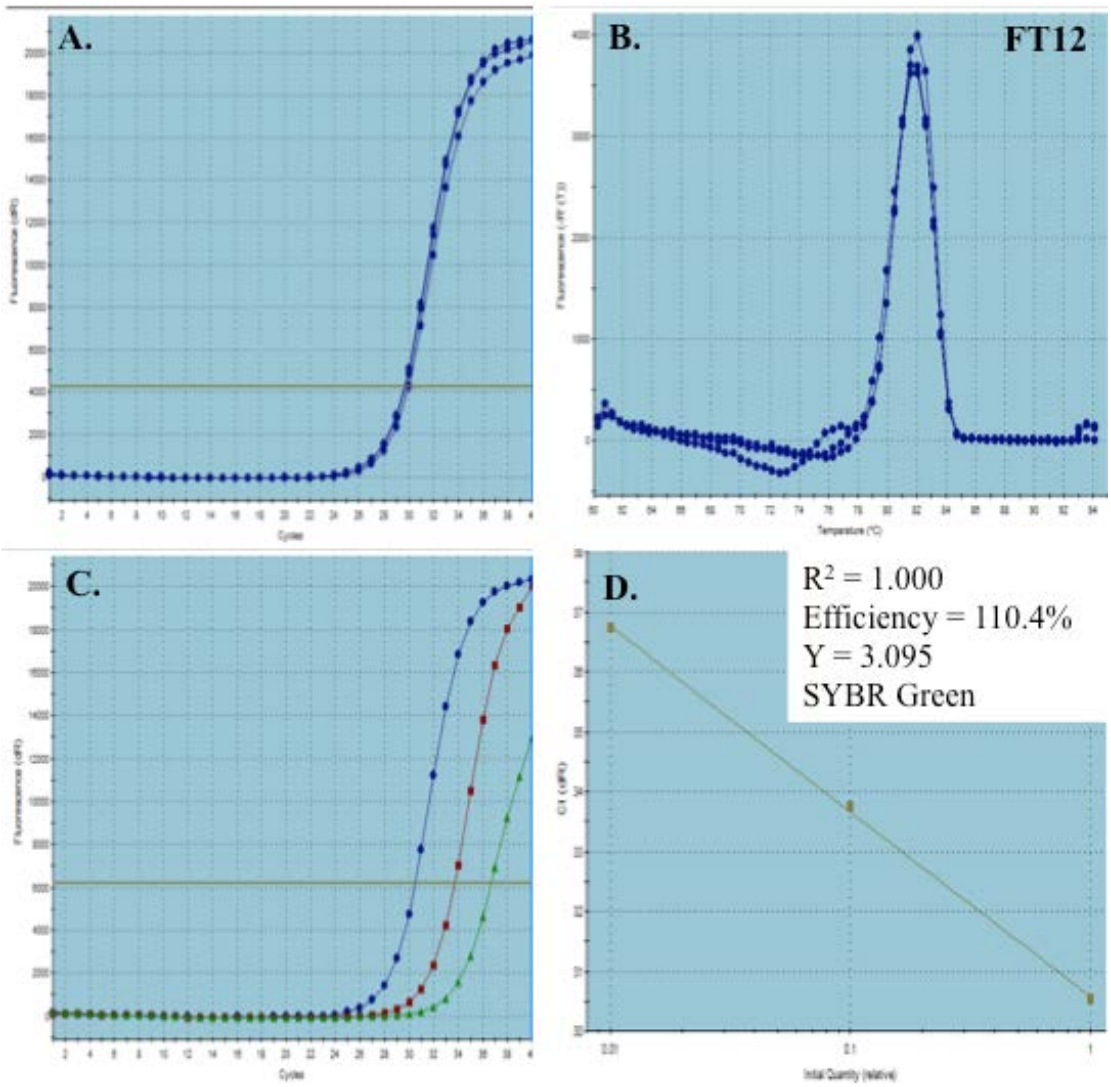

Figure S2.10. RT-qPCR graphs for primers of KfFT12 used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red) and 100x dilution (green). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 110.4%.

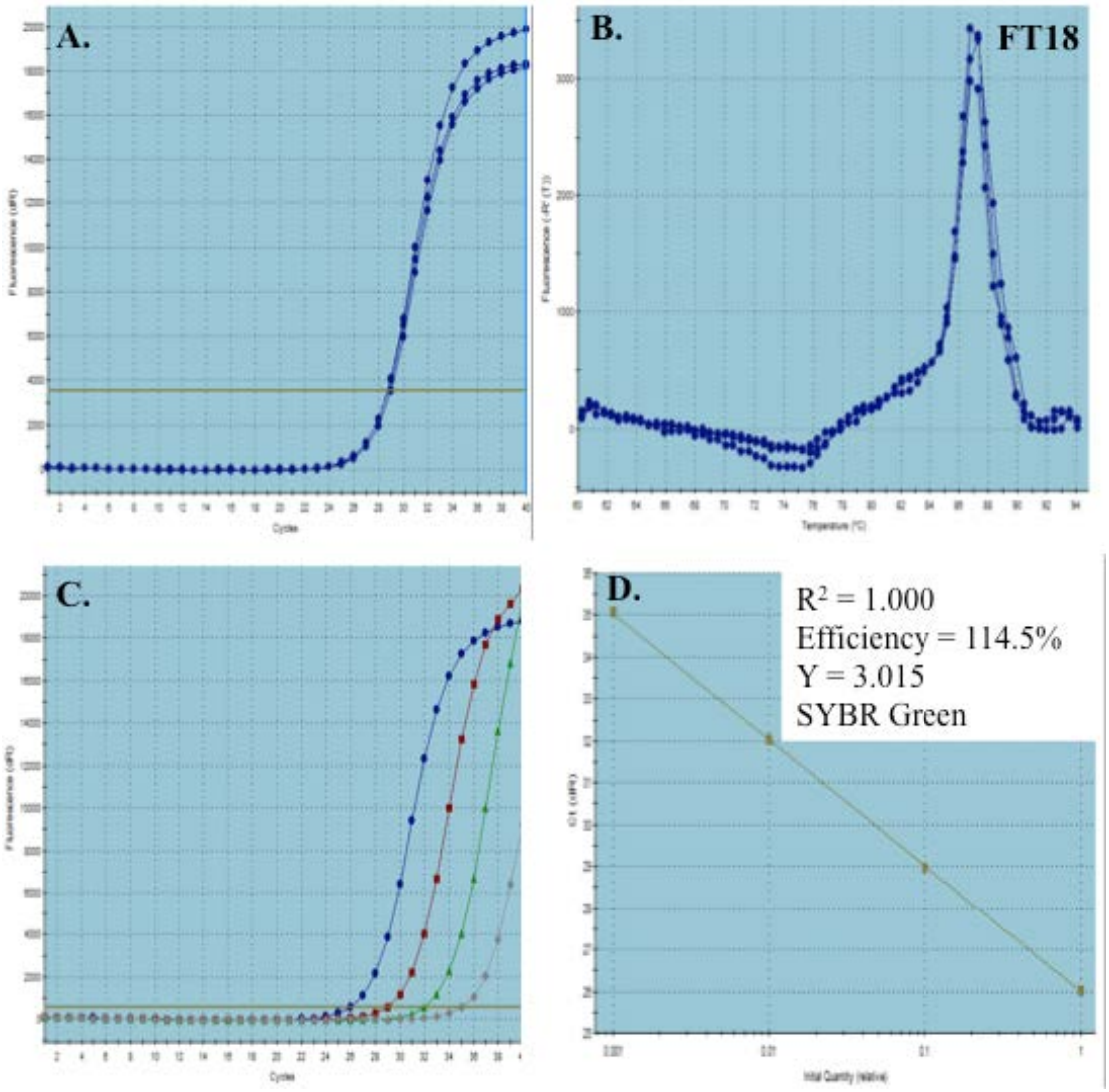

Figure S2.11. RT-qPCR graphs for primers of KfFT18 used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plot of three technical replicates of a pooled cDNA sample using 1ul of cDNA as in the actual experiment. B. Dissociation curve of cDNA at concentration used in the experiment. C. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red), 100x dilution (green) and 1000x (grey). D. Standard curve produced from using varying dilutions of cDNA in the assay mix. Efficiency of the primer set was calculated at 114.5%.

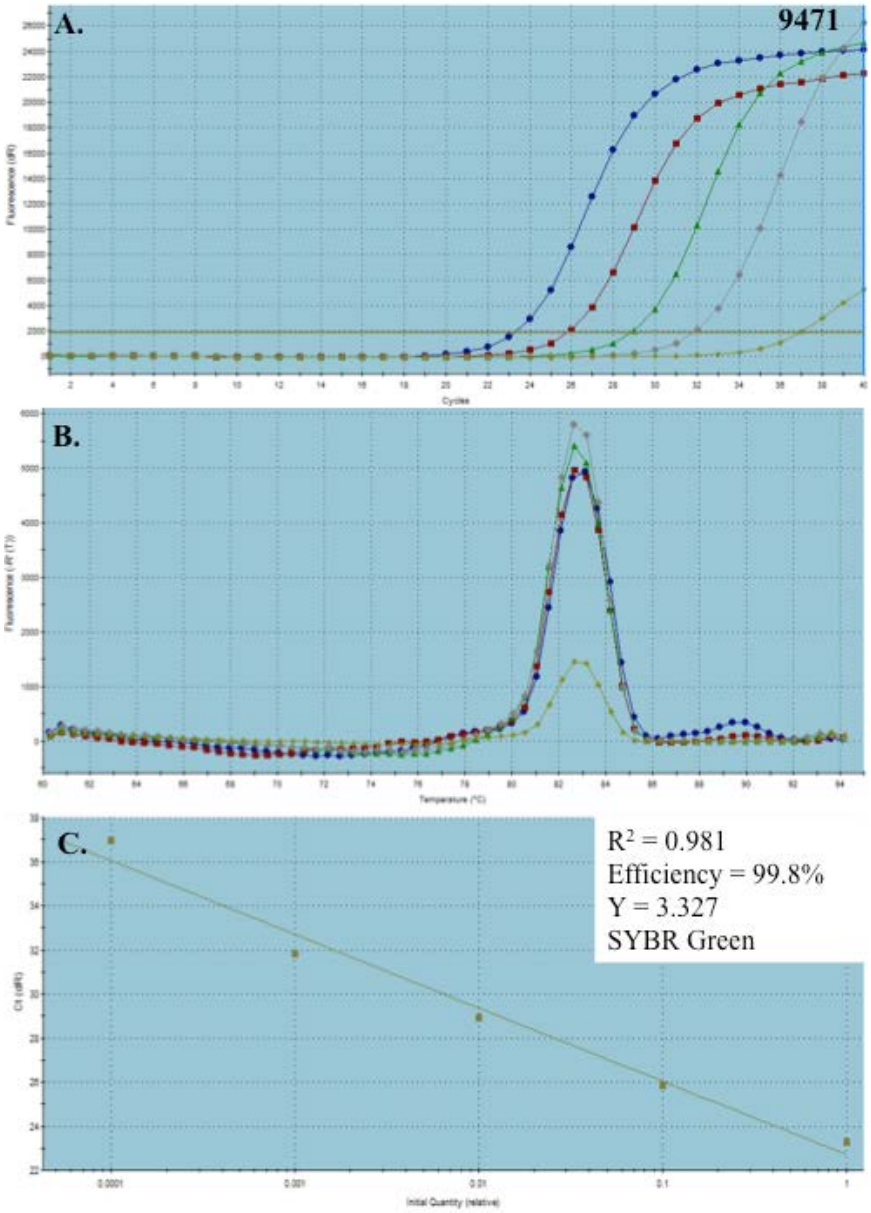

Figure S2.12. RT-qPCR graphs for primers of control gene Kf9471 used to investigate flowering initiation and regulation in *K. fedtschenkoi*. A. Amplification plots of pooled cDNA at assay concentration (blue), 10x dilution (red), 100x dilution (green), 1000x (grey) and 10000x (yellow). B. Dissociation curve of cDNA at concentration used in the experiment (blue) and dilutions explained in A. C. Standard curve produced from using varying dilutions of cDNA in the assay mix mentioned in A. Efficiency of the primer set was calculated at 99.8%.

Table S4.1. Average calculated soil moisture readings for well-watered and droughted *KfMYB439* lines over the 31day drought experiment.

Line	Watering Condition	Average Volumetric water content, Kg.L-1	Standard Error (SE)
WT	Well Watered	0.700	0.007
MYB439 FL 37C	Well Watered	0.525	0.044
MYB439 FL 54B	Well Watered	0.509	0.063
MYB439 R 3C	Well Watered	0.656	0.053
MYB439 R 10B	Well Watered	0.656	0.012
WT	Drought	0.141	0.036
MYB439 FL 37C	Drought	0.041	0.008
MYB439 FL 54B	Drought	0.035	0.006
MYB439 R 3C	Drought	0.038	0.010
MYB439 R 10B	Drought	0.028	0.003

Average capacity for well-watered plants was 0.59 $Kg.L^{-1}$ and average for the soil of droughted plants was 0.05 $Kg.L^{-1}$. Therefore with the well-watered capacity showing 100% moisture, then the droughted soil showed a vast reduction, and on average contains only 8.84% moisture.

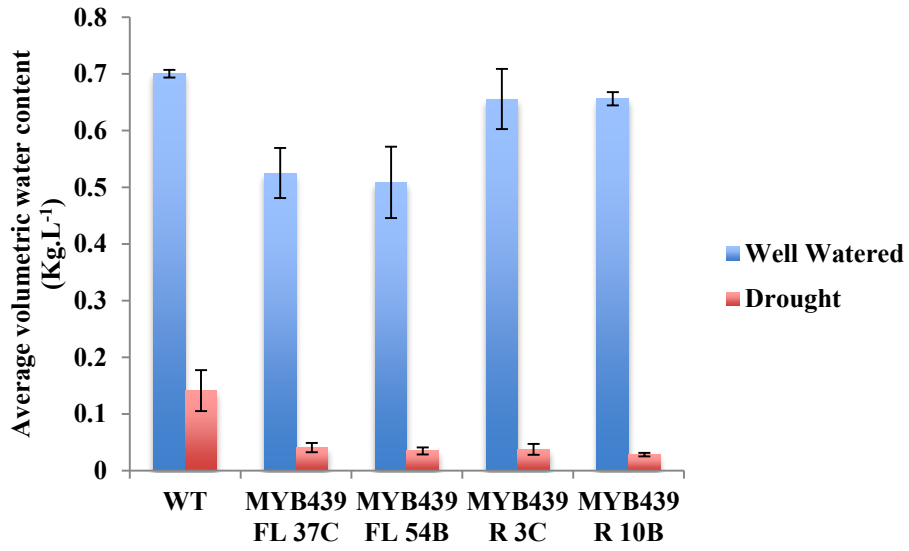

Figure S4.1. Average volumetric water content for soil over 31day time course from *KfMYB439* transgenic lines kept in well watered (blue) and droughted (red) conditions.

Chapter 5 –

Table S5.1. Average calculated soil moisture readings for well-watered and droughted *KfCIB1* lines over the 31day drought experiment.

Line	Watering Condition	Average Volumetric water content, Kg.L-1	Standard Error (SE)
WT	Well Watered	0.700	0.007
CIB FL 5B	Well Watered	0.496	0.044
CIB FL 11B	Well Watered	0.555	0.044
CIB R 12A	Well Watered	0.634	0.038
CIB R 30A	Well Watered	0.544	0.032
WT	Drought	0.141	0.036
CIB FL 5B	Drought	0.070	0.024
CIB FL 11B	Drought	0.029	0.004
CIB R 12A	Drought	0.047	0.012
CIB R 30A	Drought	0.037	0.007

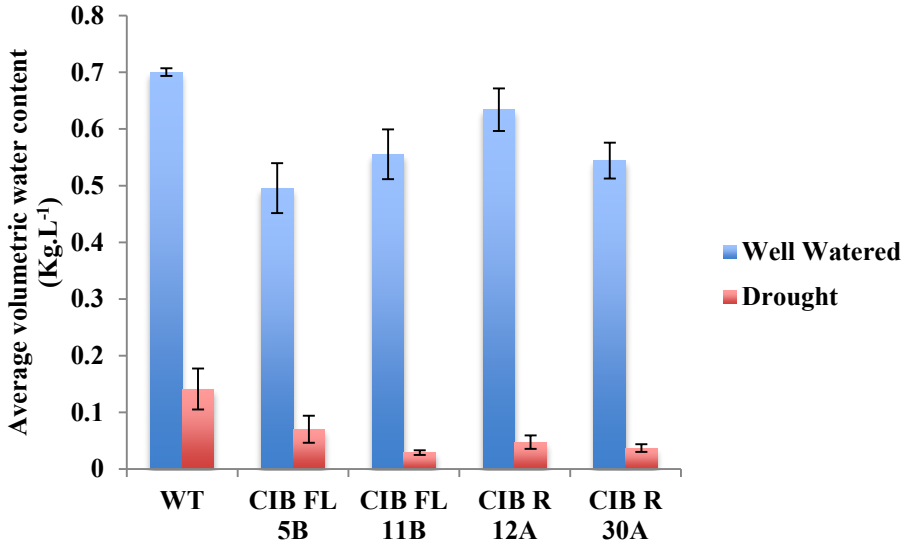

Figure S5.1. Average volumetric water content for soil over 31day time course from *KfCIB1* transgenic lines kept in well watered (blue) and droughted (red) conditions.

Table S6.1. Average calculated soil moisture readings for well-watered and droughted *KfCDF2* lines over the 31day drought experiment.

Line	Watering Condition	Average Volumetric water content, Kg.L-1	Standard Error (SE)
WT	Well Watered	0.700	0.007
CDF2 FL 13A	Well Watered	0.626	0.048
CDF2 FL 14C	Well Watered	0.649	0.015
CDF2 R 19A	Well Watered	0.680	0.025
CDF2 R 26B	Well Watered	0.527	0.051
WT	Drought	0.141	0.036
CDF2 FL 13A	Drought	0.063	0.020
CDF2 FL 14C	Drought	0.071	0.016
CDF2 R 19A	Drought	0.083	0.019
CDF2 R 26B	Drought	0.051	0.019

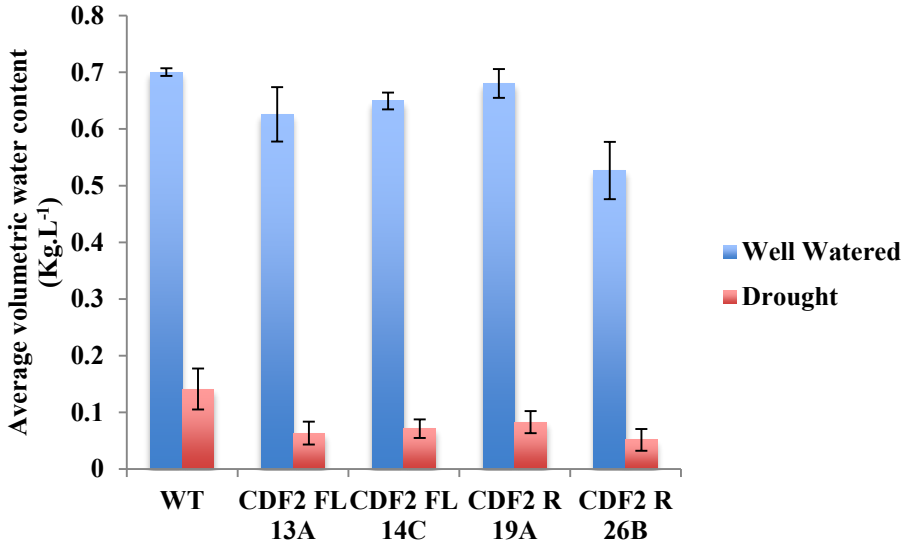

Figure S6.1. Average volumetric water content for soil over 31day time course from *KfCDF2* transgenic lines kept in well-watered (blue) and droughted (red) conditions.

Table S6.2. Average calculated soil moisture readings for well-watered and droughted *KfCDF2* lines over the 90day drought experiment.

Line	Watering Condition	Average Volumetric water content, Kg.L-1	Standard Error (SE)
WT	Well Watered	0.73	0.003
CDF2 FL 13A	Well Watered	0.726	0.021
CDF2 FL 14C	Well Watered	0.749	0.013
CDF2 R 19A	Well Watered	0.78	0.015
CDF2 R 26B	Well Watered	0.727	0.015
WT	Drought	0.061	0.021
CDF2 FL 13A	Drought	0.043	0.01
CDF2 FL 14C	Drought	0.041	0.012
CDF2 R 19A	Drought	0.043	0.02
CDF2 R 26B	Drought	0.041	0.014

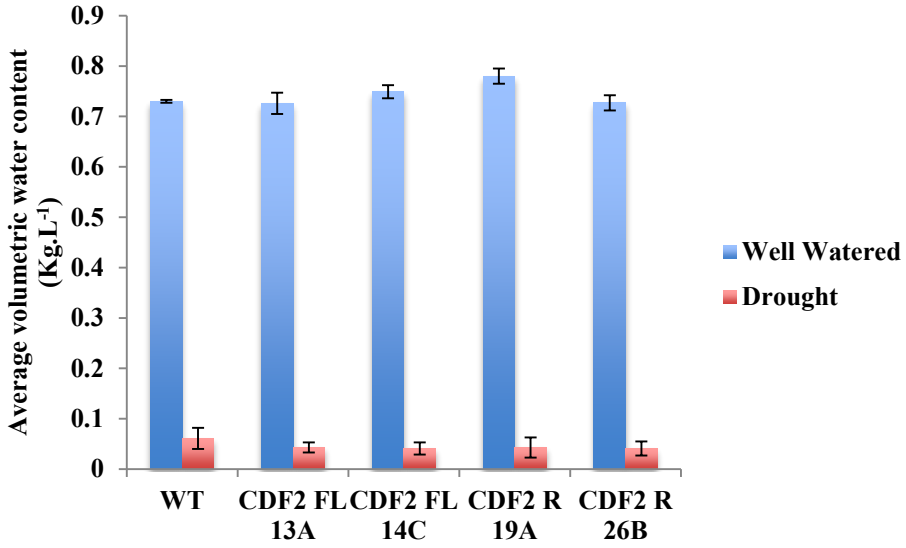

Figure S6.2. Average volumetric water content for soil over 90day time course from *KfCDF2* transgenic lines kept in well-watered (blue) and droughted (red) conditions.

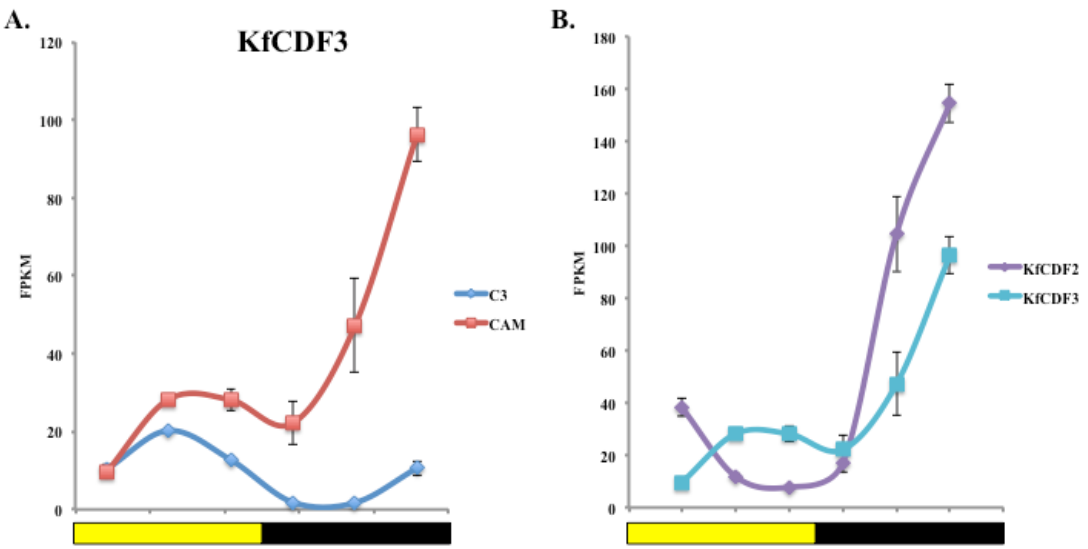

Figure S6.3. Expression profile of *KfCDF3* from SOLiD sequencing data. **A.** *KfCDF3* from SOLiD data in LP1 (C3 – blue) and LP6 (CAM – red). **B.** Comparison of expression of *KfCDF2* (purple) and *KfCDF3* (blue) in LP6 (CAM tissue).

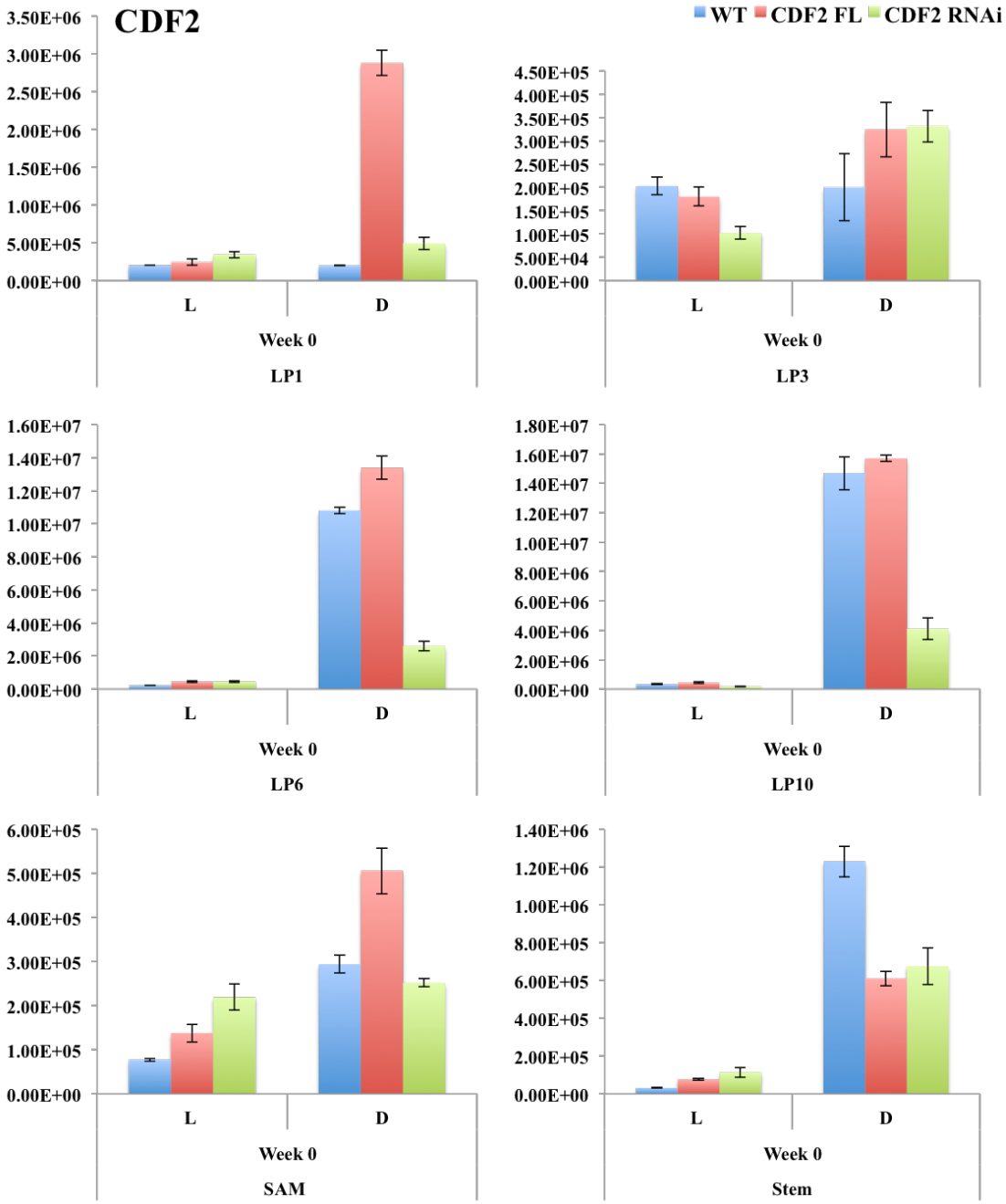

Figure S7.1. RT-qPCR data for *KfCDF2* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

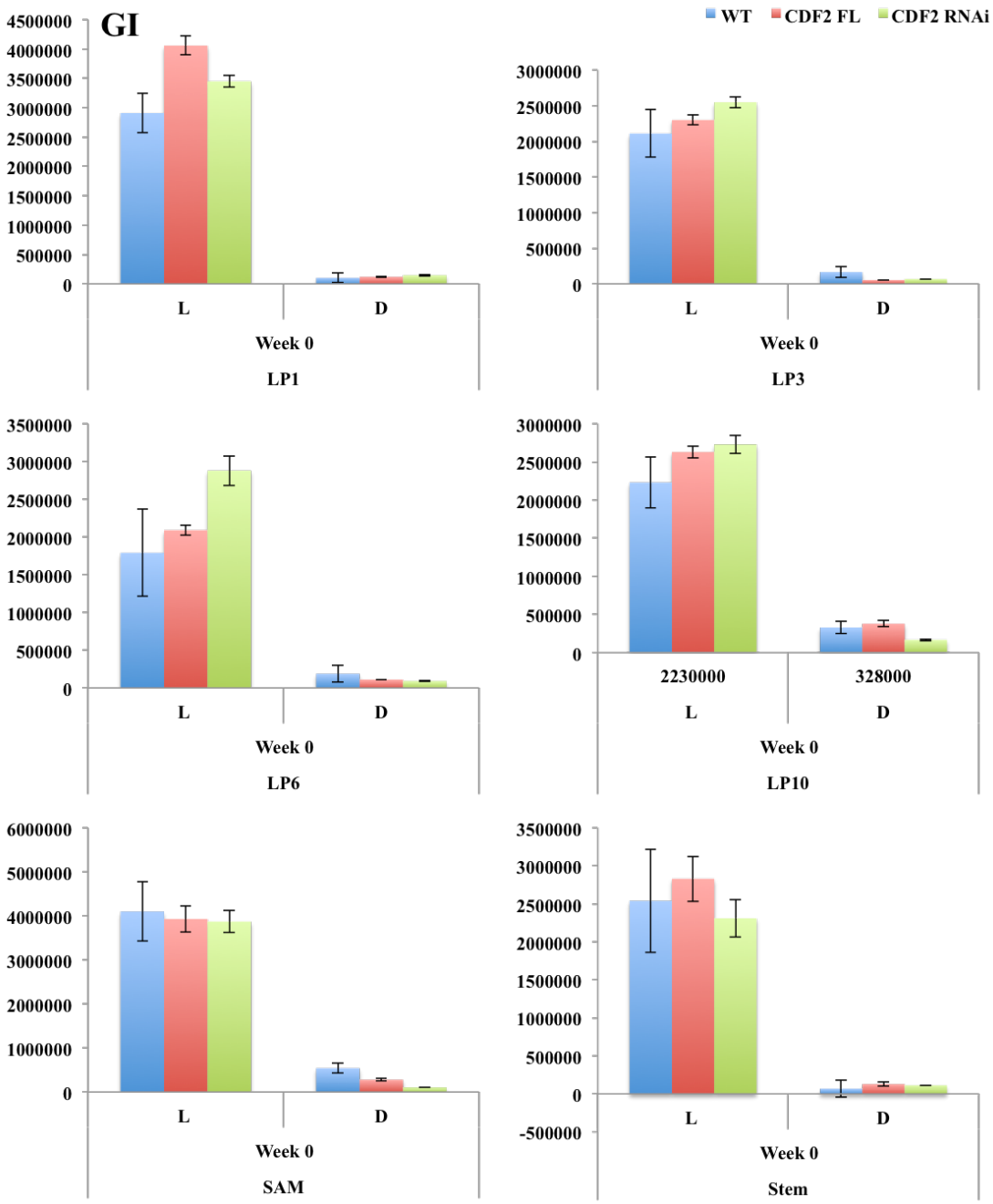

Figure S7.2. RT-qPCR data for *KfGI* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

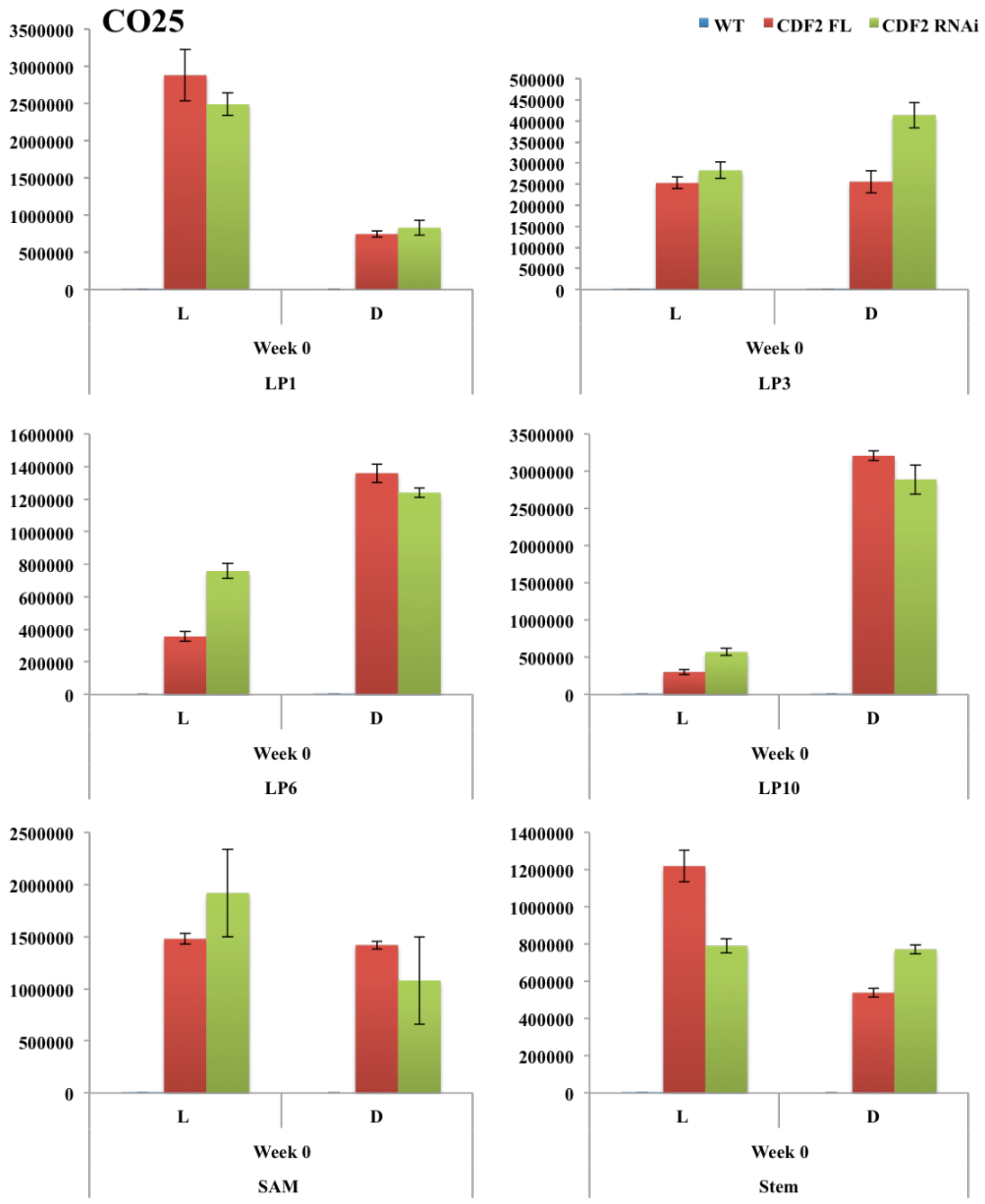

Figure S7.3. RT-qPCR data for *KfCO25* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435)

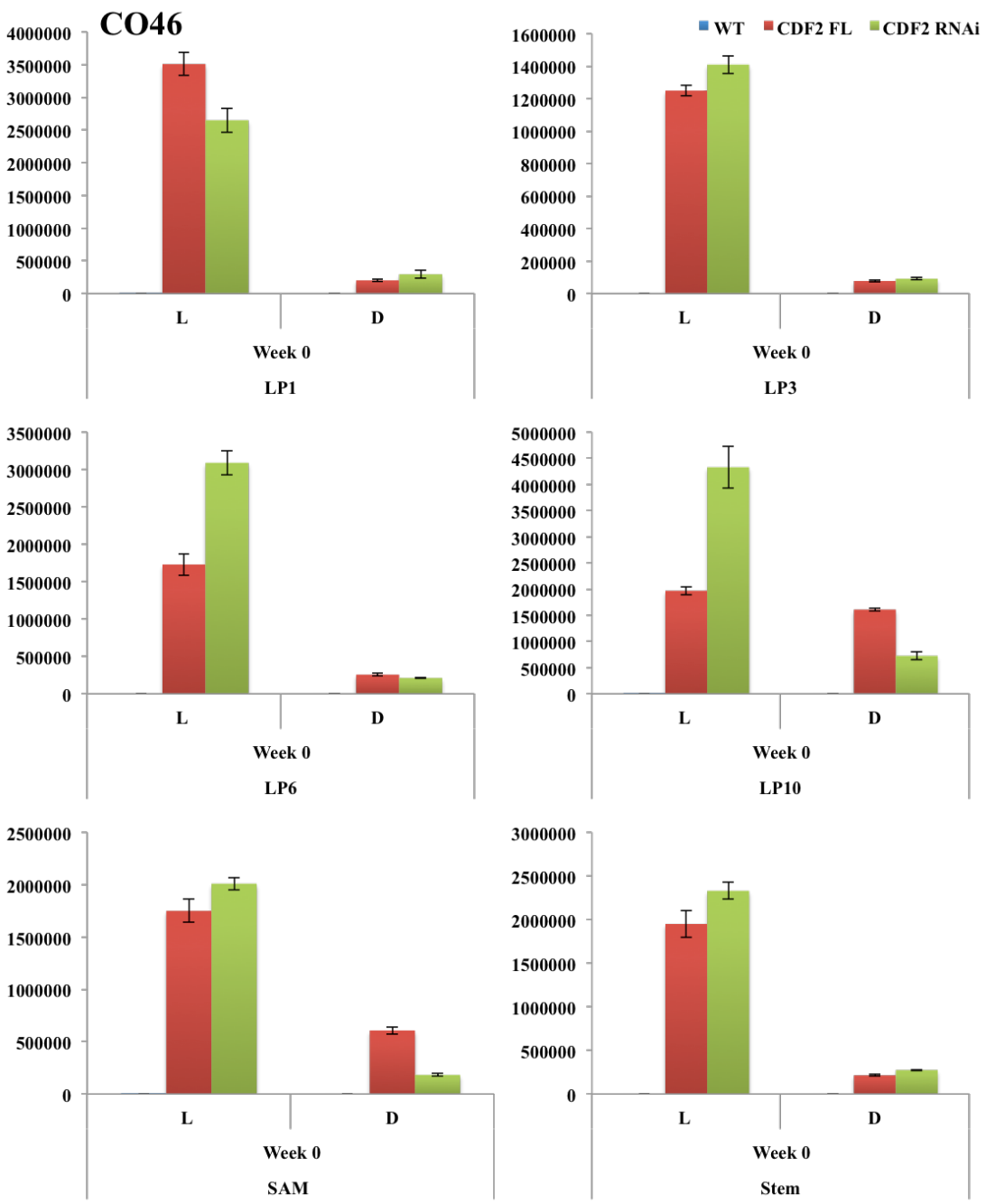

Figure S7.4. RT-qPCR data for *KfC046* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

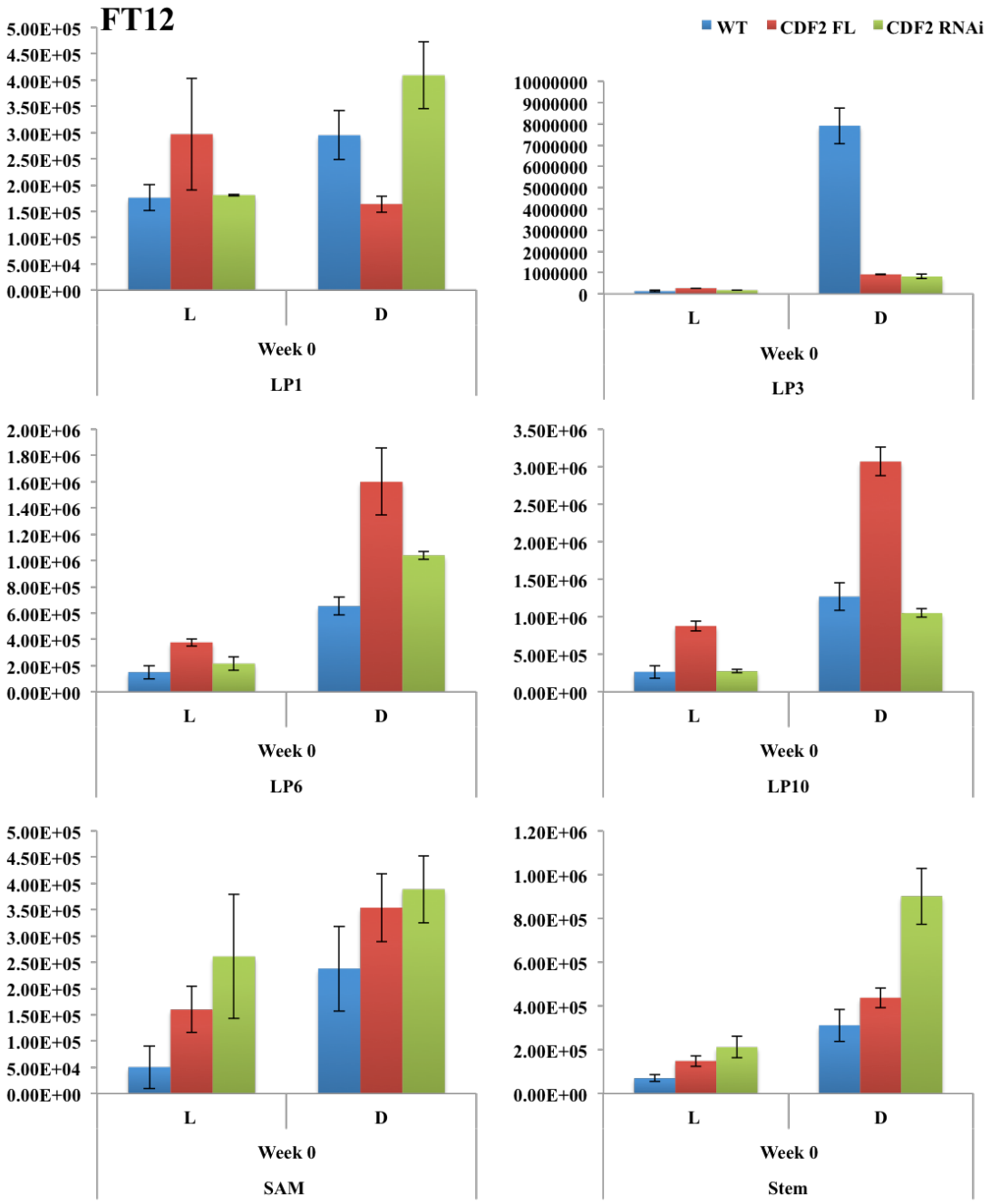

Figure S7.5. RT-qPCR data for *KfFT12* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

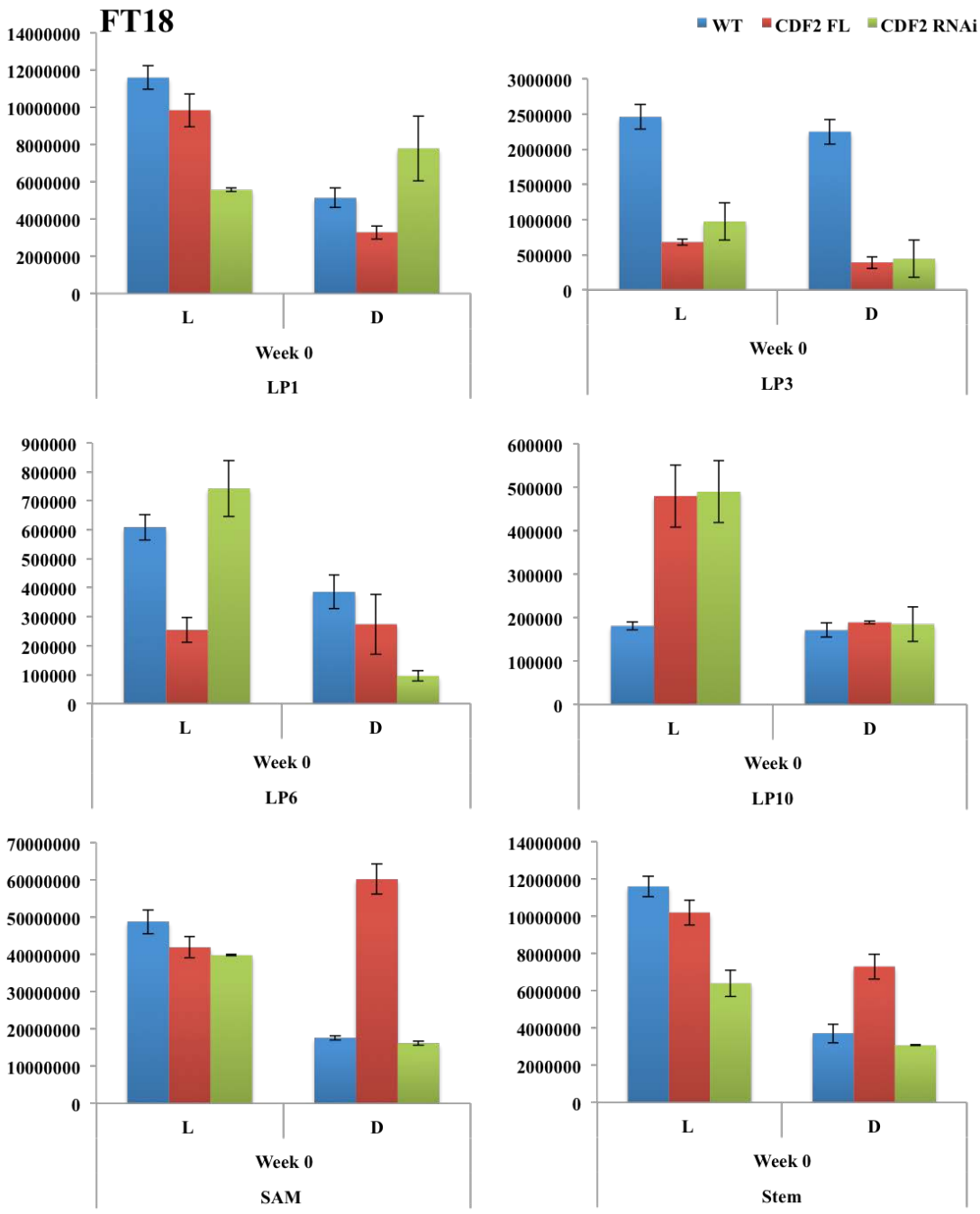

Figure S7.6. RT-qPCR data for *KfFT18* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

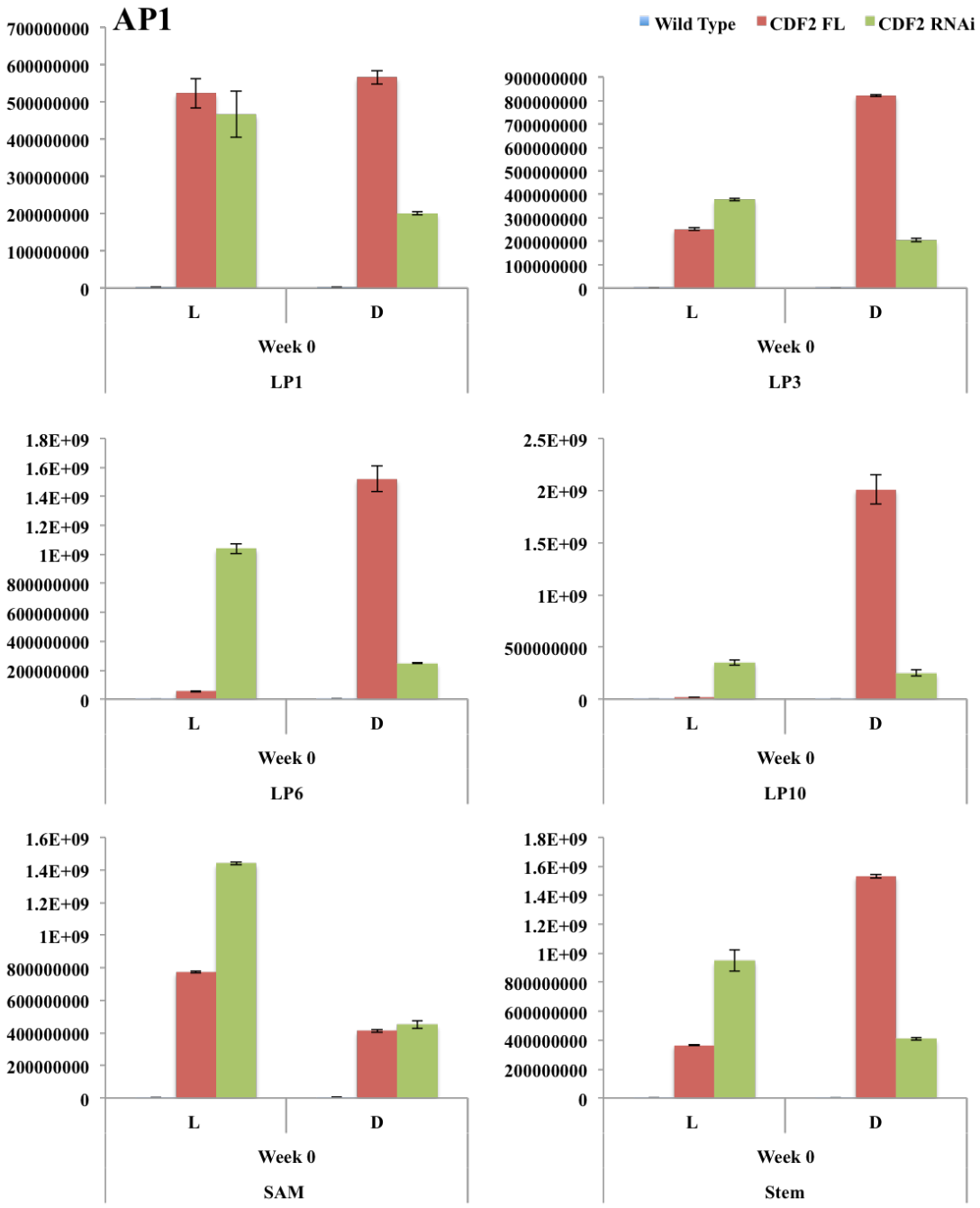

Figure S7.7. RT-qPCR data for *KfAP1* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

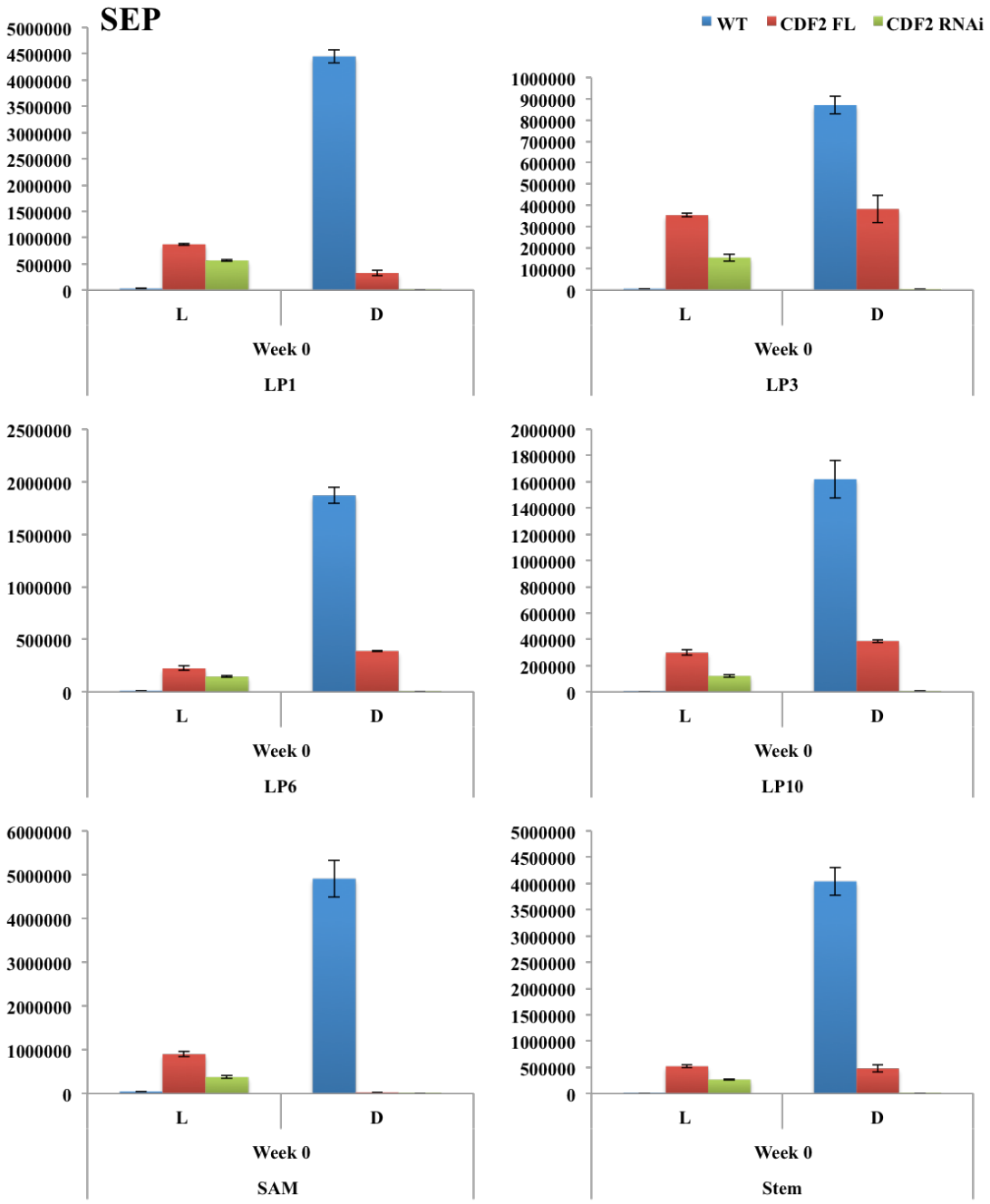

Figure S7.8. RT-qPCR data for *KfSEP* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

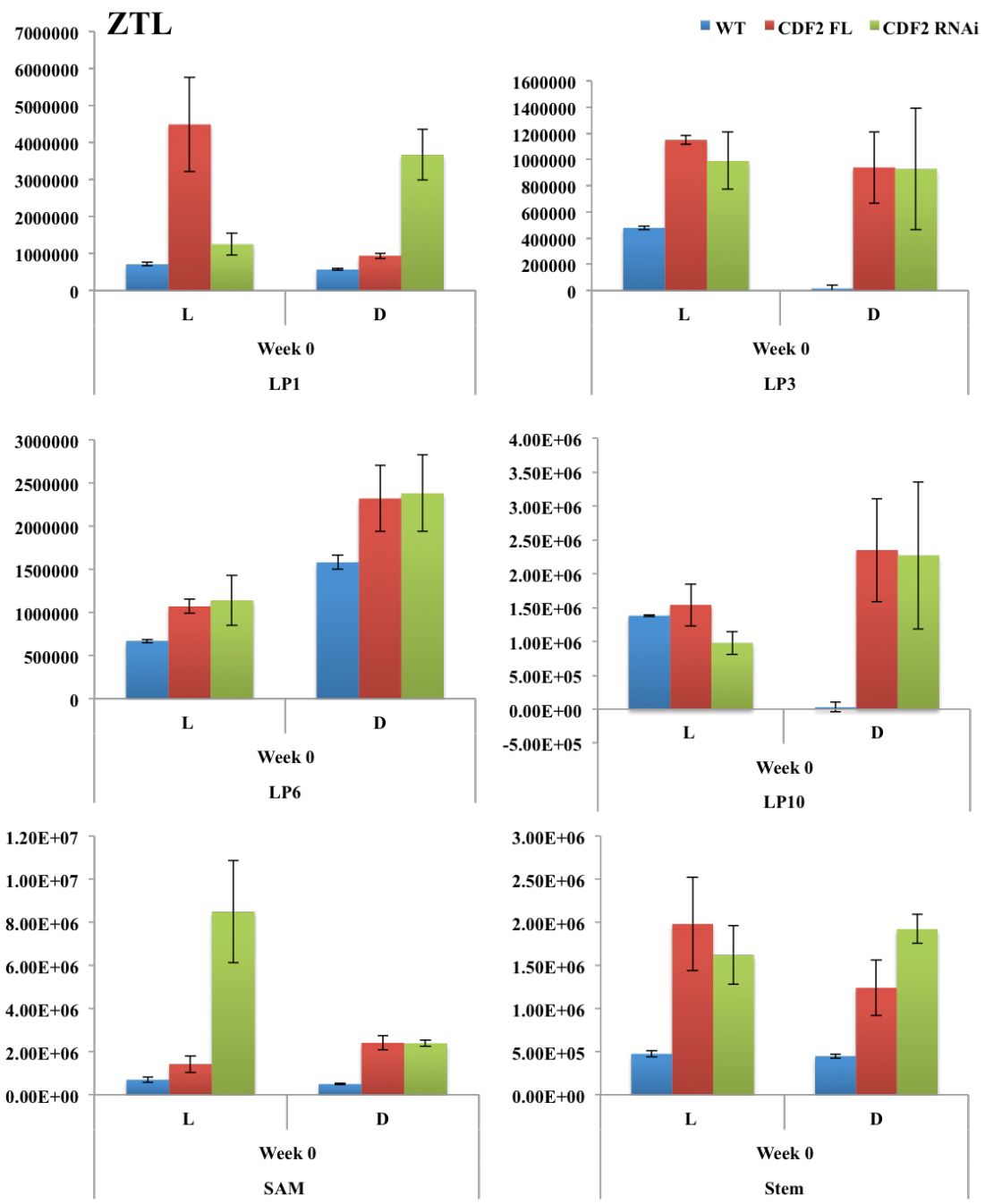

Figure S7.9. RT-qPCR data for *KfZTL* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

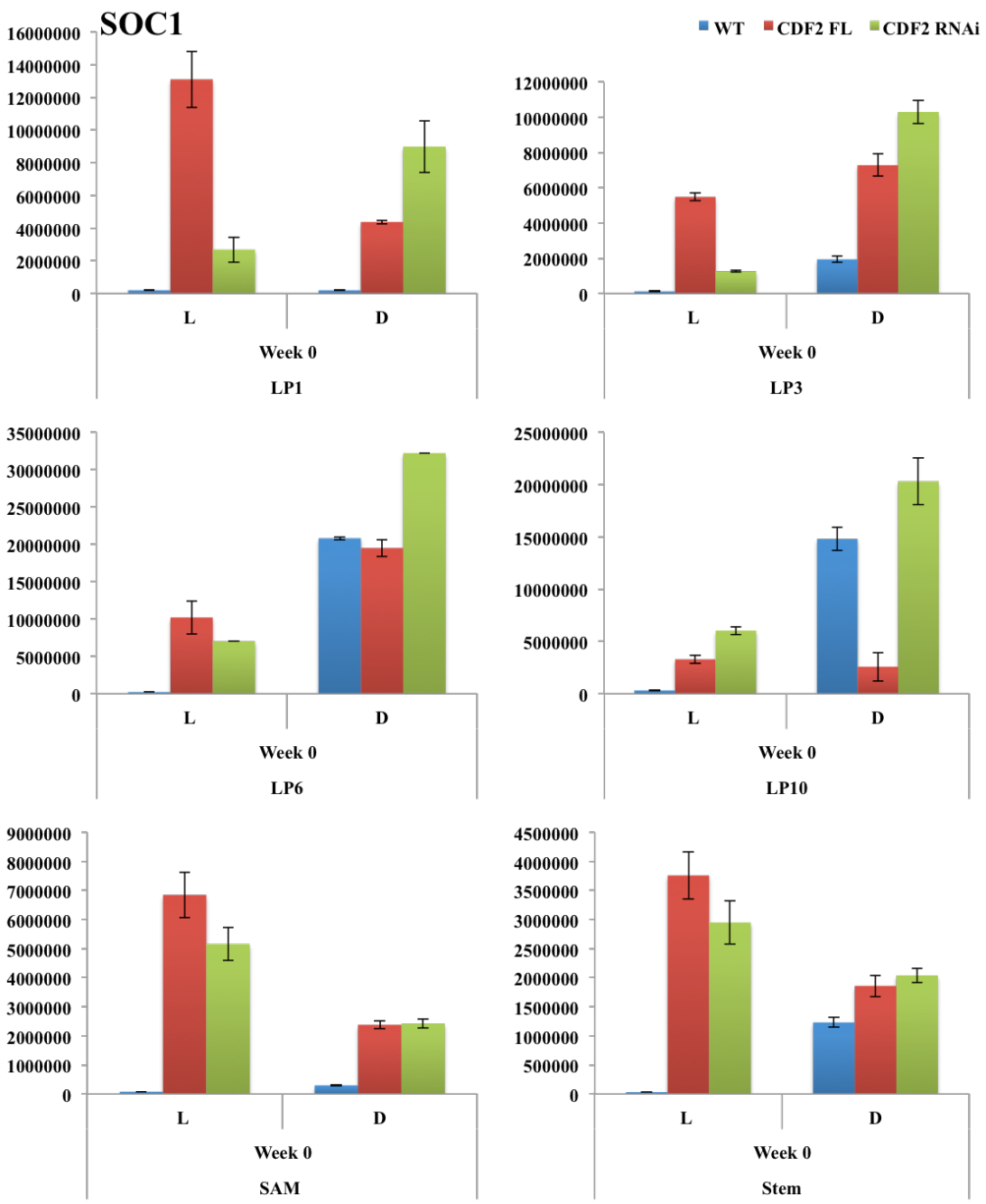

Figure S7.10. RT-qPCR data for *KfSOC1* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

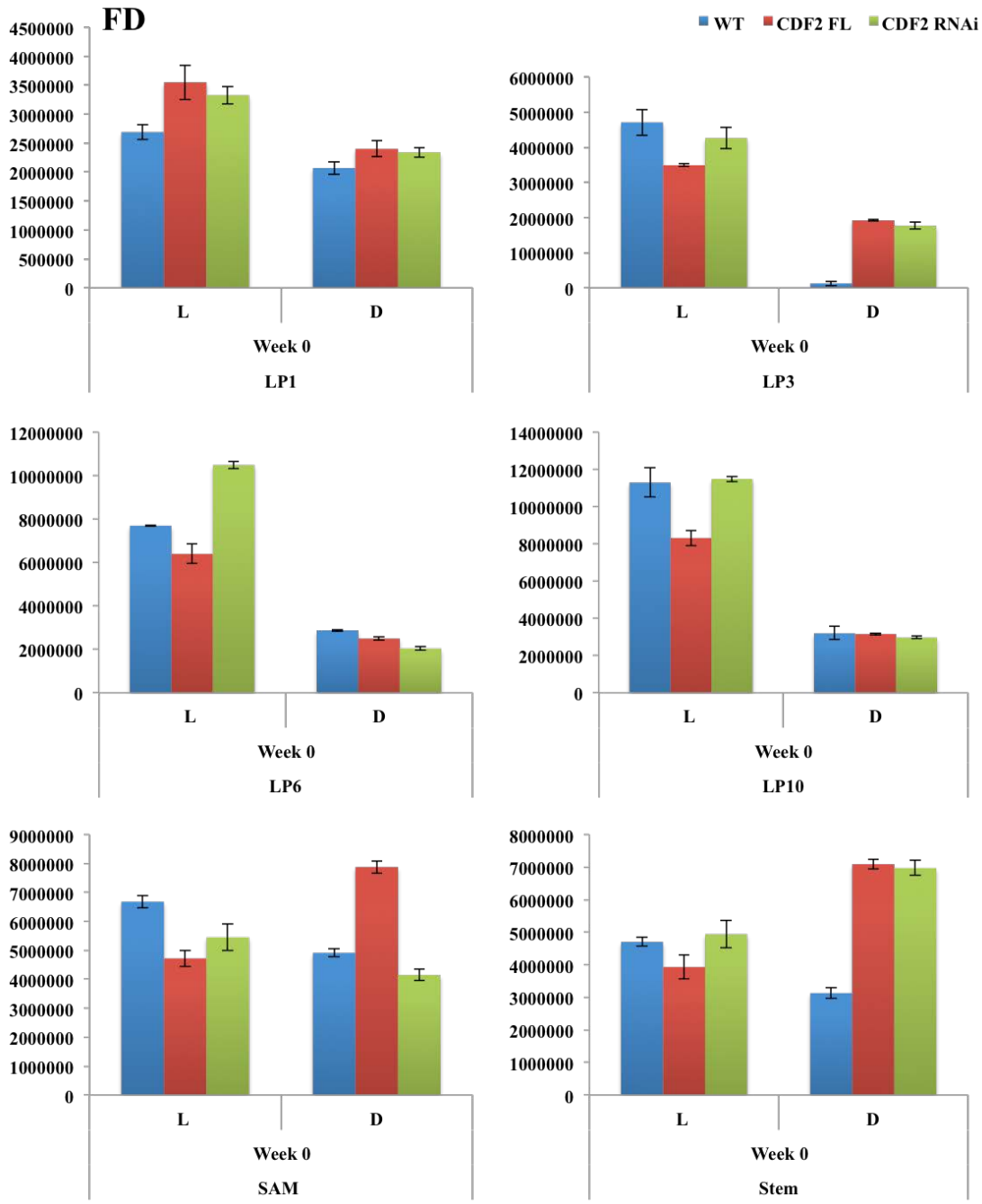

Figure S7.11. RT-qPCR data for *KfFD* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

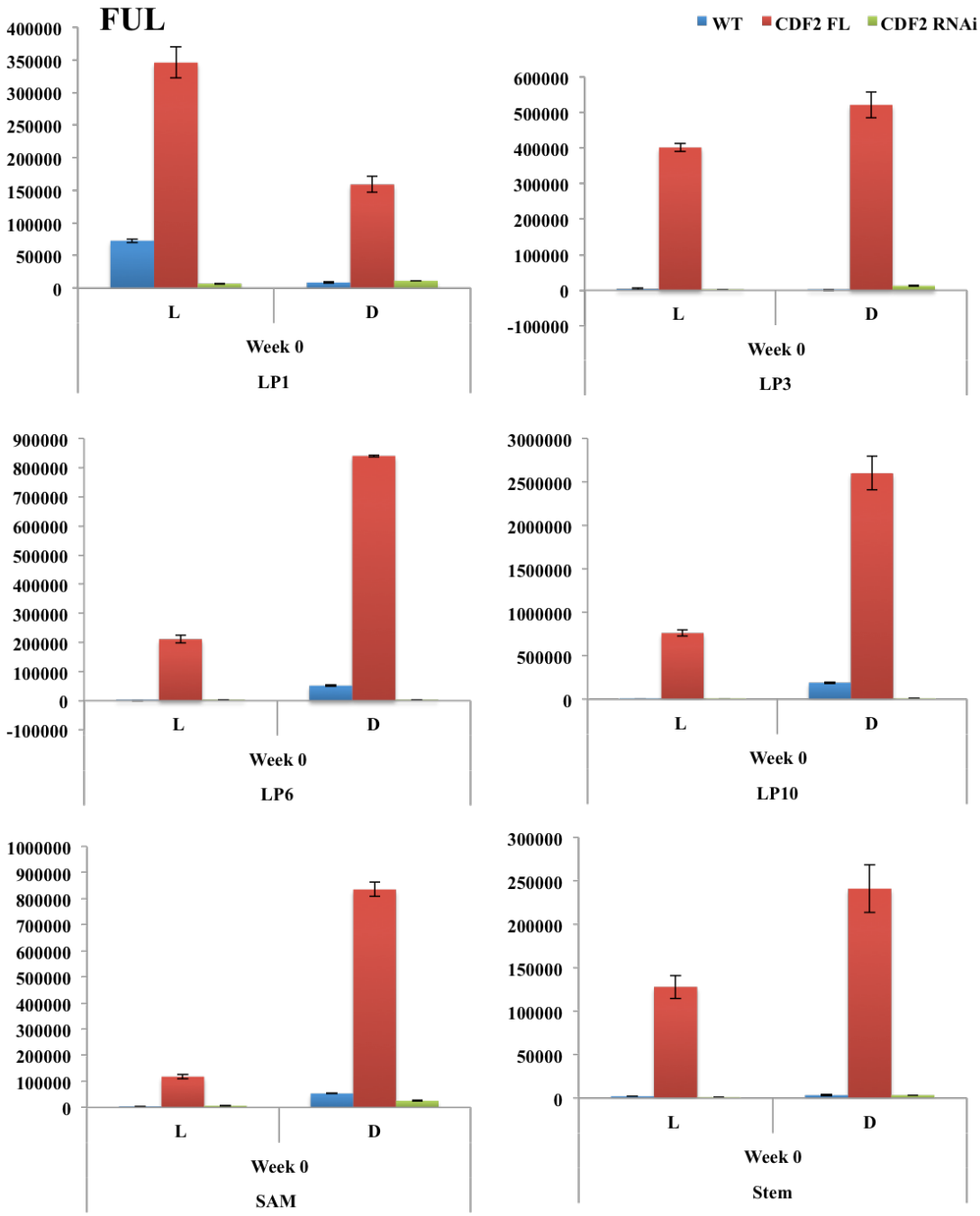

Figure S7.12. RT-qPCR data for *KfFUL* expression in wild type, *KfCDF2_FL_14C* and *KfCDF2_FL_19A* in non-flowering conditions (16:8 light/dark). Samples are: leaf paris: 1, 3, 6 and 10, shoot apical meristem, stem. L and D represent light and dark samples. Normalising gene 9471 was used (Kf gene: KF149435).

BORLAND, A. M., HARTWELL, J., JENKINS, G. I., WILKINS, M. B. & NIMMO, H. G. 1999. Metabolite control overrides circadian regulation of phosphoenolpyruvate carboxylase kinase and CO2 fixation in Crassulacean acid metabolism. *Plant Physiology,* 121**,** 889-896.

BORLAND, A. M., GRIFFITHS, H., HARTWELL, J. & SMITH, J. A. 2009. Exploiting the potential of plants with crassulacean acid metabolism for bioenergy production on marginal lands. *Journal of experimental botany,* 60**,** 2879-96.

BORLAND, A. M., HARTWELL, J., WESTON, D. J., SCHLAUCH, K. A., TSCHAPLINSKI, T. J., TUSKAN, G. A., YANG, X. & CUSHMAN, J. C. 2014. Engineering crassulacean acid metabolism to improve water-use efficiency. *Trends in plant science,* 19**,** 327-38.

BORLAND, A. M. & TAYBI, T. 2004. Synchronization of metabolic processes in plants with Crassulacean acid metabolism. *Journal of Experimental Botany,* 55**,** 1255-1265.

BORLAND, A. M., ZAMBRANO, V. A. B., CEUSTERS, J. & SHORROCK, K. 2011. The photosynthetic plasticity of crassulacean acid metabolism: an evolutionary innovation for sustainable productivity in a changing world. *New Phytologist,* 191**,** 619-633.

BOXALL, S. F., FOSTER, J. M., BOHNERT, H. J., CUSHMAN, J. C., NIMMO, H. G. & HARTWELL, J. 2005. Conservation and divergence of circadian clock operation in a stress-inducible crassulacean acid metabolism species reveals clock compensation against stress. *Plant Physiology,* 137**,** 969-982.

BUSTIN, M. 1999. Regulation of DNA-dependent a.ctivities by the functional motifs of the high-mobility-group chromosomal proteins. *Molecular and cellular biology,* 19**,** 5237-46.

BUSTIN, M., LEHN, D. A. & LANDSMAN, D. 1990. Structural features of the HMG chromosomal proteins and their genes. *Biochimica et biophysica acta,* 1049**,** 231-43.

CARLOWICZ, M. 2010. Global Temperatures. NASA Earth Observatory. Accessed 6/10/12

CARRE, I. A. & KIM, J. Y. 2002. MYB transcription factors in the Arabidopsis circadian clock. *J Exp Bot,* 53**,** 1551-7.

CARTER, P. J., NIMMO, H. G., FEWSON, C. A. & WILKINS, M. B. 1991. Circadian-Rhythms in the Activity of a Plant Protein-Kinase. *Embo Journal,* 10**,** 2063-2068.

CASPAR, T., HUBER, S. C. & SOMERVILLE, C. 1985. Alterations in Growth, Photosynthesis, and Respiration in a Starchless Mutant of Arabidopsis thaliana (L.) Deficient in Chloroplast Phosphoglucomutase Activity. *Plant physiology,* 79**,** 11-7.

CHAILAKHYAN, M. 1936. New facts in support of the hormonal theory of plant development. *CR (Doklady) Acad. Sci. URSS,* 13**,** 77.

CHALFIE, M., TU, Y., EUSKIRCHEN, G., WARD, W. W. & PRASHER, D. C. 1994. Green fluorescent protein as a marker for gene expression. *Science,* 263**,** 802-5.

CHANG, N. K. 1981. Nitrate Assimilation and Crassulacean Acid Metabolism in Leaves of Kalanchoe fedtschenkoi Variety Marginata. *Plant physiology,* 68**,** 464-8.

CHEN, W., CHAO, G. & SINGH, K. B. 1996. The promoter of a H2O2-inducible, Arabidopsis glutathione S-transferase gene contains closely linked OBF- and OBP1-binding sites. *The Plant journal : for cell and molecular biology,* 10**,** 955-66.

CHOI, H. I., HONG, J. H., HA, J. O., KANG, J. Y. & KIM, S. Y. 2000. ABFs, a family of ABA-responsive element binding factors. *Journal of Biological Chemistry,* 275**,** 1723-1730.

CHU, C., DAI, Z., KU, M. S. & EDWARDS, G. E. 1990. Induction of Crassulacean Acid Metabolism in the Facultative Halophyte Mesembryanthemum crystallinum by Abscisic Acid. *Plant Physiol,* 93**,** 1253-60.

CORBESIER, L., VINCENT, C., JANG, S., FORNARA, F., FAN, Q., SEARLE, I., GIAKOUNTIS, A., FARRONA, S., GISSOT, L., TURNBULL, C. & COUPLAND, G. 2007. FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science,* 316**,** 1030-3.

CORRALES, A. R., NEBAUER, S. G., CARRILLO, L., FERNANDEZ-NOHALES, P., MARQUES, J., RENAU-MORATA, B., GRANELL, A., POLLMANN, S., VICENTE-CARBAJOSA, J., MOLINA, R. V. & MEDINA, J. 2014. Characterization of tomato Cycling Dof Factors reveals conserved and new functions in the control of flowering time and abiotic stress responses. *Journal of experimental botany,* 65**,** 995-1012.

CORREA, L. G. G., RIANO-PACHON, D. M., SCHRAGO, C. G., DOS SANTOS, R. V., MUELLER-ROEBER, B. & VINCENTZ, M. 2008. The Role of bZIP Transcription Factors in Green Plant Evolution: Adaptive Features Emerging from Four Founder Genes. *PloS one,* 3, 32944.

COVINGTON, M. F. & HARMER, S. L. 2007. The circadian clock regulates auxin signaling and responses in Arabidopsis. *PLoS Biol,* 5**,** e222.

COVINGTON, M. F., PANDA, S., LIU, X. L., STRAYER, C. A., WAGNER, D. R. & KAY, S. A. 2001. ELF3 modulates resetting of the circadian clock in Arabidopsis. *Plant Cell,* 13**,** 1305-15.

COVSHOFF, S. & HIBBERD, J. M. 2012. Integrating C-4 photosynthesis into C-3 crops to increase yield potential. *Current Opinion in Biotechnology,* 23**,** 209-214.

CRAYN, D. M., WINTER, K. & SMITH, J. A. C. 2004. Multiple origins of crassulacean acid metabolism and the epiphytic habit in the Neotropical family Bromeliaceae. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 3703-3708.

CUSHMAN, J. C., AGARIE, S., ALBION, R. L., ELLIOT, S. M., TAYBI, T. & BORLAND, A. M. 2008a. Isolation and Characterization of Mutants of Common Ice Plant Deficient in Crassulacean Acid Metabolism. *Plant Physiol*, 147, 228-238.

CUSHMAN, J. C., TILLETT, R. L., WOOD, J. A., BRANCO, J. M. & SCHLAUCH, K. A. 2008b. Large-scale mRNA expression profiling in the common ice plant, Mesembryanthemum crystallinum, performing C3 photosynthesis and Crassulacean acid metabolism (CAM). *J. Exp. Bot.,* 59**,** 1875-1894.

DALL'OMO, C. 2011. *Timing is Everything: The molecular basis for the interaction between the circadian clock and primary metabolism un C3 and Crassulacean acid metabolism plant species.* PhD, University of Liverpool.

DE SOUZA, A. P., GASPAR, M., DA SILVA, E. A., ULIAN, E. C., WACLAWOVSKY, A. J., NISHIYAMA, M. Y., JR., DOS SANTOS, R. V., TEIXEIRA, M. M., SOUZA, G. M. & BUCKERIDGE, M. S. 2008. Elevated CO2 increases photosynthesis, biomass and productivity, and modifies gene expression in sugarcane. *Plant Cell Environ,* 31**,** 1116-27.

DEVER, L. V., BOXALL, S. F., KNEROVA, J. & HARTWELL, J. 2015. Transgenic perturbation of the decarboxylation phase of Crassulacean acid metabolism alters

physiology and metabolism but has only a small effect on growth. *Plant physiology,* 167**,** 44-59.
DEVLIN, P. F. & KAY, S. A. 2000. Cryptochromes are required for phytochrome signaling to the circadian clock but not for rhythmicity. *Plant Cell,* 12**,** 2499-2510.
DIAZ, I., VICENTE-CARBAJOSA, J., ABRAHAM, Z., MARTINEZ, M., ISABEL-LA MONEDA, I. & CARBONERO, P. 2002. The GAMYB protein from barley interacts with the DOF transcription factor BPBF and activates endosperm-specific genes during seed development. *The Plant journal : for cell and molecular biology,* 29**,** 453-64.
DING, Z. J., DOYLE, M. R., AMASINO, R. M. & DAVIS, S. J. 2007. A complex genetic interaction between Arabidopsis thaliana TOC1 and CCA1/LHY in driving the circadian clock and in output regulation. *Genetics,* 176**,** 1501-1510.
DODD, A. N., GARDNER, M. J., HOTTA, C. T., HUBBARD, K. E., DALCHAU, N., LOVE, J., ASSIE, J. M., ROBERTSON, F. C., JAKOBSEN, M. K., GONCALVES, J., SANDERS, D. & WEBB, A. A. R. 2007. The Arabidopsis circadian clock incorporates a cADPR-based feedback loop. *Science,* 318**,** 1789-1792.
DODD, A. N., GRIFFITHS, H., TAYBI, T., CUSHMAN, J. C. & BORLAND, A. M. 2003. Integrating diel starch metabolism with the circadian and environmental regulation of Crassulacean acid metabolism in Mesembryanthemum crystallinum. *Planta,* 216**,** 789-797.
DODD, A. N., PARKINSON, K. & WEBB, A. A. R. 2004. Independent circadian regulation of assimilation and stomatal conductance in the ztl-1 mutant of Arabidopsis. *New Phytologist,* 162**,** 63-70.
DODD, A. N., SALATHIA, N., HALL, A., KEVEI, E., TOTH, R., NAGY, F., HIBBERD, J. M., MILLAR, A. J. & WEBB, A. A. 2005. Plant circadian clocks increase photosynthesis, growth, survival, and competitive advantage. *Science,* 309**,** 630-3.
DOI, K., IZAWA, T., FUSE, T., YAMANOUCHI, U., KUBO, T., SHIMATANI, Z., YANO, M. & YOSHIMURA, A. 2004. Ehd1, a B-type response regulator in rice, confers short-day promotion of flowering and controls FT-Iike gene expression independently of Hd1l. *Genes & Development,* 18**,** 926-936.
DUBOS, C., STRACKE, R., GROTEWOLD, E., WEISSHAAR, B., MARTIN, C. & LEPINIEC, L. 2010. MYB transcription factors in Arabidopsis. *Trends Plant Sci,* 15**,** 573-81.
DUCAT, D. C. & SILVER, P. A. 2012. Improving carbon fixation pathways. *Curr Opin Chem Biol,* 16**,** 337-344.
ECKEL-MAHAN, K. & SASSONE-CORSI, P. 2013. Metabolism and the circadian clock converge. *Physiol Rev,* 93**,** 107-35.
EHLERT, A., WELTMEIER, F., WANG, X., MAYER, C. S., SMEEKENS, S., VICENTE-CARBAJOSA, J. & DROGE-LASER, W. 2006. Two-hybrid protein-protein interaction analysis in Arabidopsis protoplasts: establishment of a heterodimerization map of group C and group S bZIP transcription factors. *The Plant journal : for cell and molecular biology,* 46**,** 890-900.
FARINAS, B. & MAS, P. 2011. Functional implication of the MYB transcription factor RVE8/LCL5 in the circadian control of histone acetylation. *The Plant journal : for cell and molecular biology,* 66**,** 318-29.
FARRE, E. M., HARMER, S. L., HARMON, F. G., YANOVSKY, M. J. & KAY, S. A. 2005. Overlapping and distinct roles of PRR7 and PRR9 in the Arabidopsis circadian clock. *Current Biology,* 15**,** 47-54.

FELDBRUGGE, M., SPRENGER, M., HAHLBROCK, K. & WEISSHAAR, B. 1997. PcMYB1, a novel plant protein containing a DNA-binding domain with one MYB repeat, interacts in vivo with a light-regulatory promoter unit. *The Plant journal : for cell and molecular biology,* 11**,** 1079-93.

FIELDS, S. & SONG, O. 1989. A novel genetic system to detect protein-protein interactions. *Nature,* 340**,** 245-6.

FILICHKIN, S. A., DIFAZIO, S. P., BRUNNER, A. M., DAVIS, J. M., YANG, Z. K., KALLURI, U. C., ARIAS, R. S., ETHERINGTON, E., TUSKAN, G. A. & STRAUSS, S. H. 2007. Efficiency of gene silencing repeats vs. transitive RNAi in Arabidopsis: direct inverted vectors. *Plant Biotechnology Journal,* 5**,** 615-626.

FILICHKIN, S. A., PRIEST, H. D., GIVAN, S. A., SHEN, R., BRYANT, D. W., FOX, S. E., WONG, W. K. & MOCKLER, T. C. 2010. Genome-wide mapping of alternative splicing in Arabidopsis thaliana. *Genome Res,* 20**,** 45-58.

FORNARA, F., DE MONTAIGU, A., SANCHEZ-VILLARREAL, A., TAKAHASHI, Y., VER LOREN VAN THEMAAT, E., HUETTEL, B., DAVIS, S. J. & COUPLAND, G. 2015. The GI-CDF module of Arabidopsis affects freezing tolerance and growth as well as flowering. *The Plant journal : for cell and molecular biology,* 81**,** 695-706.

FORNARA, F., PANIGRAHI, K. C., GISSOT, L., SAUERBRUNN, N., RUHL, M., JARILLO, J. A. & COUPLAND, G. 2009. Arabidopsis DOF transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. *Dev Cell,* 17**,** 75-86.

FOSTER, R., IZAWA, T. & CHUA, N. H. 1994. Plant Bzip Proteins Gather at Acgt Elements. *Faseb J,* 8**,** 192-200.

FUKAZAWA, J., SAKAI, T., ISHIDA, S., YAMAGUCHI, I., KAMIYA, Y. & TAKAHASHI, Y. 2000. REPRESSION OFSHOOT GROWTH, a bZIP transcriptional activator, regulates cell elongation by controlling the level of gibberellins. *Plant Cell,* 12**,** 901-915.

GABRIELE, S., RIZZA, A., MARTONE, J., CIRCELLI, P., COSTANTINO, P. & VITTORIOSO, P. 2010. The Dof protein DAG1 mediates PIL5 activity on seed germination by negatively regulating GA biosynthetic gene AtGA3ox1. *The Plant journal : for cell and molecular biology,* 61**,** 312-23.

GALIMBA, K. D. & DI STILIO, V. S. 2015. Sub-functionalization to ovule development following duplication of a floral organ identity gene. *Dev Biol,* 405**,** 158-72.

GARCES, H. M., CHAMPAGNE, C. E., TOWNSLEY, B. T., PARK, S., MALHO, R., PEDROSO, M. C., HARADA, J. J. & SINHA, N. R. 2007. Evolution of asexual reproduction in leaves of the genus Kalanchoe. *Proceedings of the National Academy of Sciences of the United States of America,* 104**,** 15578-83.

GENDRON, J. M., PRUNEDA-PAZ, J. L., DOHERTY, C. J., GROSS, A. M., KANG, S. E. & KAY, S. A. 2012. Arabidopsis circadian clock protein, TOC1, is a DNA-binding transcription factor. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 3167-3172.

GIBALOVA, A., RENAK, D., MATCZUK, K., DUPL'AKOVA, N., CHAB, D., TWELL, D. & HONYS, D. 2009. AtbZIP34 is required for Arabidopsis pollen wall patterning and the control of several metabolic pathways in developing pollen. *Plant Molecular Biology,* 70**,** 581-601.

GOULD, P. D., LOCKE, J. C. W., LARUE, C., SOUTHERN, M. M., DAVIS, S. J., HANANO, S., MOYLE, R., MILICH, R., PUTTERILL, J., MILLAR, A. J. & HALL, A. 2006. The molecular basis of temperature compensation in the Arabidopsis circadian clock. *Plant Cell,* 18**,** 1177-1187.

GRAF, A., SCHLERETH, A., STITT, M. & SMITH, A. M. 2010. Circadian control of carbohydrate availability for growth in Arabidopsis plants at night. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 20, 9458-63.

GRIFFITHS, H., HELLIKER, B., ROBERTS, A., HASLAM, R. P., GIRNUS, J., ROBE, W. E., BORLAND, A. M. & MAXWELL, K. 2002. Regulation of Rubisco activity in crassulacean acid metabolism plants: better late than never. *Functional Plant Biology,* 29**,** 689-696.

GUALBERTI, G., PAPI, M., BELLUCCI, L., RICCI, I., BOUCHEZ, D., CAMILLERI, C., COSTANTINO, P. & VITTORIOSO, P. 2002. Mutations in the Dof zinc finger genes DAG2 and DAG1 influence with opposite effects the germination of Arabidopsis seeds. *Plant Cell,* 14**,** 1253-63.

GUTIERREZ, R. A., STOKES, T. L., THUM, K., XU, X., OBERTELLO, M., KATARI, M. S., TANURDZIC, M., DEAN, A., NERO, D. C., MCCLUNG, C. R. & CORUZZI, G. M. 2008. Systems approach identifies an organic nitrogen-responsive gene network that is regulated by the master clock control gene CCA1. *Proc Natl Acad Sci U S A,* 105**,** 4939-44.

HAIDER, M. S., BARNES, J. D., CUSHMAN, J. C. & BORLAND, A. M. 2012. A CAM- and starch-deficient mutant of the facultative CAM species Mesembryanthemum crystallinum reconciles sink demands by repartitioning carbon during acclimation to salinity. *J Exp Bot,* 63**,** 1985-96.

HARMER, S. L. 2009. The Circadian System in Higher Plants. *Annual Review of Plant Biology,* 60.

HANANO, S., DOMAGALSKA, M. A., NAGY, F. & DAVIS, S. J. 2006. Multiple phytohormones influence distinct parameters of the plant circadian clock. *Genes to Cells,* 11**,** 1381-1392.

HARTWELL, J. 2005. The co-ordination of central plant metabolism by the circadian clock. *Biochemical Society Transactions,* 33**,** 945-948.

HARTWELL, J. 2006. The circadian clock in CAM plants. *In:* HALL, A. J. W. & MCWATTERS, H. G. (eds.) *Annual plant reviews: endogenous plant rhythms.* Oxford, UK: Blackwell Publishing.

HARTWELL, J., GILL, A., NIMMO, G. A., WILKINS, M. B., JENKINS, G. I. & NIMMO, H. G. 1999. Phosphoenolpyruvate carboxylase kinase is a novel protein kinase regulated at the level of expression. *The Plant Journal* 20.

HARTWELL, J., SMITH, L. H., WILKINS, M. B., JENKINS, G. I. & NIMMO, H. G. 1996. Higher plant phosphoenolpyruvate carboxylase kinase is regulated at the level of translatable mRNA in response to light or a circadian rhythm. *Plant Journal,* 10**,** 1071-1078.

HAYAMA, R., YOKOI, S., TAMAKI, S., YANO, M. & SHIMAMOTO, K. 2003. Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature,* 422**,** 719-722.

HAYDON, M. J., MIELCZAREK, O., ROBERTSON, F. C., HUBBARD, K. E. & WEBB, A. A. 2013. Photosynthetic entrainment of the Arabidopsis thaliana circadian clock. *Nature,* 502**,** 689-92.

HERNANDO-AMADO, S., GONZALEZ-CALLE, V., CARBONERO, P. & BARRERO-SICILIA, C. 2012. The family of DOF transcription factors in Brachypodium distachyon: phylogenetic comparison with rice and barley DOFs and expression profiling. *BMC Plant Biol,* 12**,** 202.

HIGUCHI, Y., NARUMI, T., ODA, A., NAKANO, Y., SUMITOMO, K., FUKAI, S. & HISAMATSU, T. 2013. The gated induction system of a systemic floral inhibitor,

antiflorigen, determines obligate short-day flowering in chrysanthemums. *Proceedings of the National Academy of Sciences of the United States of America,* 110**,** 17137-42.
HOCHULI, P. A. & FEIST-BURKHARDT, S. 2013. Angiosperm-like pollen and Afropollis from the Middle Triassic (Anisian) of the Germanic Basin (Northern Switzerland). *Frontiers in plant science,* 4**,** 344.
HSU, P. Y., DEVISETTY, U. K. & HARMER, S. L. 2013. Accurate timekeeping is controlled by a cycling activator in Arabidopsis. *eLife,* 2**,** e00473.
HUANG, T., BOHLENIUS, H., ERIKSSON, S., PARCY, F. & NILSSON, O. 2005. The mRNA of the Arabidopsis gene FT moves from leaf to shoot apex and induces flowering. *Science,* 309**,** 1694-6.
HURST, H. C. 1995. Transcription factors 1: bZIP proteins. *Protein profile,* 2**,** 101-68.
HWANG, M. G., CHUNG, I. K., KANG, B. G. & CHO, M. H. 2001. Sequence-specific binding property of Arabidopsis thaliana telomeric DNA binding protein 1 (AtTBP1). *FEBS Lett,* 503**,** 35-40.
IMAIZUMI, T., SCHULTZ, T. F., HARMON, F. G., HO, L. A. & KAY, S. A. 2005. FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in Arabidopsis. *Science,* 309**,** 293-7.
IZAWA, T., FOSTER, R. & CHUA, N. H. 1993. Plant Bzip Protein-DNA Binding-Specificity. *J Mol Biol,* 230**,** 1131-1144.
IZAWA, T., OIKAWA, T., SUGIYAMA, N., TANISAKA, T., YANO, M. & SHIMAMOTO, K. 2002. Phytochrome mediates the external light signal to repress FT orthologs in photoperiodic flowering of rice. *Genes & Development,* 16**,** 2006-2020.
JAKOBY, M., WEISSHAAR, B., DROGE-LASER, W., VICENTE-CARBAJOSA, J., TIEDEMANN, J., KROJ, T., PARCY, F. & GRP, B. R. 2002. bZIP transcription factors in Arabidopsis. *Trends in Plant Science,* 7**,** 106-111.
JAMES, A. B., SYED, N. H., BORDAGE, S., MARSHALL, J., NIMMO, G. A., JENKINS, G. I., HERZYK, P., BROWN, J. W. & NIMMO, H. G. 2012. Alternative splicing mediates responses of the Arabidopsis circadian clock to temperature changes. *Plant Cell,* 24**,** 961-81.
JIAO, Y. L., LAU, O. S. & DENG, X. W. 2007. Light-regulated transcriptional networks in higher plants. *Nature Reviews Genetics,* 8**,** 217-230.
JONES, M. B. 1975. The effect of leaf age on leaf resistance and CO2 exchange of the CAM plant Bryophyllum fedtschenkoi. *Planta,* 123**,** 91-6.
KANG, H. G. & SINGH, K. B. 2000. Characterization of salicylic acid-responsive, arabidopsis Dof domain proteins: overexpression of OBP3 leads to growth defects. *The Plant journal : for cell and molecular biology,* 21**,** 329-39.
KEVEI, E., GYULA, P., FEHER, B., TOTH, R., VICZIAN, A., KIRCHER, S., REA, D., DORJGOTOV, D., SCHAFER, E., MILLAR, A. J., KOZMA-BOGNAR, L. & NAGY, F. 2007. Arabidopsis thaliana circadian clock is regulated by the small GTPase LIP1. *Curr Biol,* 17**,** 1456-64.
KAUFMANN, K., MUINO, J. M., OSTERAS, M., FARINELLI, L., KRAJEWSKI, P. & ANGENENT, G. C. 2010. Chromatin immunoprecipitation (ChIP) of plant transcription factors followed by sequencing (ChIP-SEQ) or hybridization to whole genome arrays (ChIP-CHIP). *Nature protocols,* 5**,** 457-72.
KIMURA, Y., AOKI, S., ANDO, E., KITATSUJI, A., WATANABE, A., OHNISHI, M., TAKAHASHI, K., INOUE, S., NAKAMICHI, N., TAMADA, Y. & KINOSHITA, T. 2015. A flowering integrator, SOC1, affects stomatal opening in Arabidopsis thaliana. *Plant Cell Physiol,* 56**,** 640-9.

KISAKA, H., YANAGISAWA, S., MIWA, T. & AKIYAMA, A. 2007. Potatoes having an increased yield of starch per plant body and method for producing the same. Google Patents.
KISU, Y., ONO, T., SHIMOFURUTANI, N., SUZUKI, M. & ESAKA, M. 1998. Characterization and expression of a new class of zinc finger protein that binds to silencer region of ascorbate oxidase gene. *Plant and Cell Physiology,* 39**,** 1054-1064.
KLUGE, M. & BRULFERT, J. 1996. Crassulacean acid metabolism in the genus Kalanchoe: Ecological, physiological and biochemical aspects. *In:* WINTER, K. & SMITH, J. A. C. (eds.) *Crassulacean Acid Metabolism - Biochemistry, Ecophysiology and Evolution.*
KNOTT, J. E. 1934. Effect of localized photoperiod on spinach. *Proc. Soc. Hort. Sci.,* 31**,** 152-154.
KOLB, C. A., KASER, M. A., KOPECKY, J., ZOTZ, G., RIEDERER, M. & PFUNDEL, E. E. 2001. Effects of natural intensities of visible and ultraviolet radiation on epidermal ultraviolet screening and photosynthesis in grape leaves. *Plant Physiol,* 127**,** 863-75.
KOMEDA, Y. 2004. Genetic regulation of time to flower in Arabidopsis thaliana. *Annual Review of Plant Biology,* 55**,** 521-535.
KROHN, N. M., YANAGISAWA, S. & GRASSER, K. D. 2002. Specificity of the stimulatory interaction between chromosomal HMGB proteins and the transcription factor Dof2 and its negative regulation by protein kinase CK2-mediated phosphorylation. *The Journal of biological chemistry,* 277**,** 32438-44.
KROUK, G., GUTIERREZ, R. A., GIFFORD, M. G., BIRNBAUM, K. D., SHASHA, G. E. & CORUZZI, G. 2008. A systems approach to nitrogen regulatory networks and the Virtual Plant. *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology,* 150**,** S46-S46.
LEE, B. J., PARK, C. J., KIM, S. K., KIM, K. J. & PAEK, K. H. 2006. In vivo binding of hot pepper bZIP transcription factor CabZIP1 to the G-box region of pathogenesis-related protein 1 promoter. *Biochemical and biophysical research communications,* 344**,** 55-62.
LEE, J. & LEE, I. 2010. Regulation and function of SOC1, a flowering pathway integrator. *Journal of Experimental Botany,* 61**,** 2247-2254.
LEEGOOD, R. C. 2013. Strategies for engineering C-4 photosynthesis. *J. Plant Physiol.,* 170**,** 378-388.
LIAO, Y., ZOU, H. F., WANG, H. W., ZHANG, W. K., MA, B., ZHANG, J. S. & CHEN, S. Y. 2008. Soybean GmMYB76, GmMYB92, and GmMYB177 genes confer stress tolerance in transgenic Arabidopsis plants. *Cell research,* 18**,** 1047-60.
LIU, J., CHEN, N., CHEN, F., CAI, B., DAL SANTO, S., TORNIELLI, G. B., PEZZOTTI, M. & CHENG, Z. M. 2014. Genome-wide analysis and expression profile of the bZIP transcription factor gene family in grapevine (Vitis vinifera). *BMC genomics,* 15**,** 281.
LIU, L., LIU, C., HOU, X., XI, W., SHEN, L., TAO, Z., WANG, Y. & YU, H. 2012. FTIP1 is an essential regulator required for florigen transport. *PLoS Biol,* 10**,** e1001313.
LU, C. A., HO, T. H., HO, S. L. & YU, S. M. 2002. Three novel MYB proteins with one DNA binding repeat mediate sugar and hormone regulation of alpha-amylase gene expression. *Plant Cell,* 14**,** 1963-80.
LUTTGE, U. 2002. CO2-concentrating: consequences in crassulacean acid metabolism.

LUTTGE, U. 2004. Ecophysiology of Crassulacean Acid Metabolism (CAM). *Annals of Botany,* 93**,** 629-652.
LUTTGE, U. & BECK, F. 1992. Endogenous rhythms and chaos in crassulacean acid metabolism. *Planta,* 188**,** 28-38.
MAIER, T., GUELL, M. & SERRANO, L. 2009. Correlation of mRNA and protein in complex biological samples. *FEBS Lett,* 583**,** 3966-73.
MANDEL, M. A., GUSTAFSON-BROWN, C., SAVIDGE, B. & YANOFSKY, M. F. 1992. Molecular characterization of the Arabidopsis floral homeotic gene APETALA1. *Nature,* 360**,** 273-7.
MARQUEZ, Y., BROWN, J. W., SIMPSON, C., BARTA, A. & KALYNA, M. 2012. Transcriptome survey reveals increased complexity of the alternative splicing landscape in Arabidopsis. *Genome Res,* 22**,** 1184-95.
MARTIN, C. & PAZ-ARES, J. 1997. MYB transcription factors in plants. *Trends in genetics : TIG,* 13**,** 67-73.
MARTIN-TRYON, E. L. & HARMER, S. L. 2008. XAP5 CIRCADIAN TIMEKEEPER coordinates light signals for proper timing of photomorphogenesis and the circadian clock in Arabidopsis. *Plant Cell,* 20**,** 1244-59.
MARTIN-TRYON, E. L., KREPS, J. A. & HARMER, S. L. 2007. GIGANTEA acts in blue light signaling and has biochemically separable roles in circadian clock and flowering time regulation. *Plant Physiology,* 143**,** 473-486.
MARTINOIA, E., MAESHIMA, M. & NEUHAUS, H. E. 2007. Vacuolar transporters and their essential role in plant metabolism. Symposium of the Society-for-Experimental-Biology held in Honour of Georges Bernier, Apr Jan 2006 Canterbury, ENGLAND. OXFORD: Oxford Univ Press, 83-102.
MAURINO, V. G. & WEBER, A. P. M. 2013. Engineering photosynthesis in plants and synthetic microorganisms. *Journal of Experimental Botany,* 64**,** 743-751.
MAS, P., KIM, W. Y., SOMERS, D. E. & KAY, S. A. 2003. Targeted degradation of TOC1 by ZTL modulates circadian function in Arabidopsis thaliana. *Nature,* 426**,** 567-570.
MATHIEU, J., WARTHMANN, N., KUTTNER, F. & SCHMID, M. 2007. Export of FT protein from phloem companion cells is sufficient for floral induction in Arabidopsis. *Curr Biol,* 17**,** 1055-60.
MAXWELL, K., BORLAND, A. M., HASLAM, R. P., HELLIKER, B. R., ROBERTS, A. & GRIFFITHS, H. 1999. Modulation of Rubisco activity during the diurnal phases of the Crassulacean acid metabolism plant Kalanchoe daigremontiana. *Plant Physiology,* 121**,** 849-856.
MCWATTERS, H. G., BASTOW, R. M., HALL, A. & MILLAR, A. J. 2000. The ELF3 zeitnehmer regulates light signalling to the circadian clock. *Nature,* 408**,** 716-20.
MCWATTERS, H. G., KOLMOS, E., HALL, A., DOYLE, M. R., AMASINO, R. M., GYULA, P., NAGY, F., MILLAR, A. J. & DAVIS, S. J. 2007. ELF4 is required for oscillatory properties of the circadian clock. *Plant Physiol,* 144**,** 391-401.
MESHI, T. & IWABUCHI, M. 1995. Plant transcription factors. *Plant Cell Physiol,* 36**,** 1405-20.
MESZTER, R. 2010. *Molecular characterisation of transcription factors with potential roles in the circadian regulation of Crassulacean acid metabolism in Kalanchoë fedtschenkoi.* PhD, University of Liverpool.
MESZTER, R., BOXALL, S., JONES, M. & HARTWELL, J. 2008. Molecular dissection of the circadian clock output pathway that mediates the temporal coordination of Crassulacean acid metabolism. *Comparative Biochemistry and Physiology a-Molecular & Integrative Physiology,* 150**,** S153-S153.

MIAO, Y., JIANG, J., REN, Y. & ZHAO, Z. 2013. The single-stranded DNA-binding protein WHIRLY1 represses WRKY53 expression and delays leaf senescence in a developmental stage-dependent manner in Arabidopsis. *Plant physiology,* 163**,** 746-56.

MICHAEL, T. P. & MCCLUNG, C. R. 2002. Phase-specific circadian clock regulatory elements in Arabidopsis. *Plant Physiol,* 130**,** 627-38.

MIZOGUCHI, T., WHEATLEY, K., HANZAWA, Y., WRIGHT, L., MIZOGUCHI, M., SONG, H. R., CARRE, I. A. & COUPLAND, G. 2002. LHY and CCA1 are partially redundant genes required to maintain circadian rhythms in Arabidopsis. *Dev Cell,* 2**,** 629-41.

MIZOGUCHI, T., WRIGHT, L., FUJIWARA, S., CREMER, F., LEE, K., ONOUCHI, H., MOURADOV, A., FOWLER, S., KAMADA, H., PUTTERILL, J. & COUPLAND, G. 2005. Distinct roles of GIGANTEA in promoting flowering and regulating circadian rhythms in Arabidopsis. *Plant Cell,* 17**,** 2255-70.

MOLLERING, H. 1974. L-Malate: determination with malate dehydrogenase and gllutamate-oxaloacetate transaminase. *Methods of enzymatic analysis.* Verlan Chemie, Weinheim, Germany: Bergmeyer HU.

MOYANO, E., MARTINEZGARCIA, J. F. & MARTIN, C. 1996. Apparent redundancy in Myb gene function provides gearing for the control of flavonoid biosynthesis in Antirrhinum flowers. *Plant Cell,* 8**,** 1519-1532.

NAKAJIMA, K., SENA, G., NAWY, T. & BENFEY, P. N. 2001. Intercellular movement of the putative transcription factor SHR in root patterning. *Nature,* 413**,** 307-11.

NEGI, J., MORIWAKI, K., KONISHI, M., YOKOYAMA, R., NAKANO, T., KUSUMI, K., HASHIMOTO-SUGIMOTO, M., SCHROEDER, J. I., NISHITANI, K., YANAGISAWA, S. & IBA, K. 2013. A Dof transcription factor, SCAP1, is essential for the development of functional stomata in Arabidopsis. *Current biology : CB,* 23**,** 479-84.

NIMMO, G. A., NIMMO, H. G., HAMILTON, I. D., FEWSON, C. A. & WILKINS, M. B. 1986. Purification of the phosphorylated night form and dephosphorylated day form of phosphoenolpyruvate carboxylase from Bryophyllum fedtschenkoi. *The Biochemical journal,* 239**,** 213-20.

NIMMO, G. A., NIMMO, H. G., FEWSON, C. A. & WILKINS, M. B. 1984. Diurnal changes in the properties of phosphenolpyruvate carboxylase in Bryophyllum leaves: a possible covalent modification. *FEBS Lett,* 178**,** 199-203.

NIMMO, G. A., WILKINS, M. B., FEWSON, C. A. & NIMMO, H. G. 1987. Persistent circadian rhythms in the phosphorylation state of phosphoenolpyruvate carboxylase from Bryophyllum fedtschenkoi leaves and in its sensitivity to inhibition by malate. *Planta,* 170**,** 408-15.

NOBEL, P. S. 1996. High productivity of certain agronomic CAM species. *In:* WINTER, K. & SMITH, J. A. C. (eds.) *Crassulacean Acid Metabolism - Biochemistry, Ecophysiology and Evolution.*

NUSINOW, D. A., HELFER, A., HAMILTON, E. E., KING, J. J., IMAIZUMI, T., SCHULTZ, T. F., FARRE, E. M. & KAY, S. A. 2011. The ELF4-ELF3-LUX complex links the circadian clock to diurnal control of hypocotyl growth. *Nature,* 475**,** 398-402.

O'LEARY, M., H. 1982. Phosphoenolpyruvate carboxylase: an enzymologist's view. *Annu Rev Plant Physiol,* 33**,** 297-315.

O'MAOILEIDIGH, D. S., GRACIET, E. & WELLMER, F. 2014. Gene networks controlling Arabidopsis thaliana flower development. *New Phytologist,* 201**,** 16-30.

PALTIEL, J., AMIN, R., GOVER, A., ORI, N. & SAMACH, A. 2006. Novel roles for GIGANTEA revealed under environmental conditions that modify its expression in Arabidopsis and Medicago truncatula. *Planta,* 224**,** 1255-68.

PAPI, M., SABATINI, S., ALTAMURA, M. M., HENNIG, L., SCHAFER, E., COSTANTINO, P. & VITTORIOSO, P. 2002. Inactivation of the phloem-specific Dof zinc finger gene DAG1 affects response to light and integrity of the testa of Arabidopsis seeds. *Plant Physiol,* 128**,** 411-7.

PAPI, M., SABATINI, S., BOUCHEZ, D., CAMILLERI, C., COSTANTINO, P. & VITTORIOSO, P. 2000. Identification and disruption of an Arabidopsis zinc finger gene controlling seed germination. *Genes & Development,* 14**,** 28-33.

PARK, D. H., LIM, P. O., KIM, J. S., CHO, D. S., HONG, S. H. & NAM, H. G. 2003. The Arabidopsis COG1 gene encodes a Dof domain transcription factor and negatively regulates phytochrome signaling. *The Plant journal : for cell and molecular biology,* 34**,** 161-71.

PARRY, M. L., CANZIANI, O. F., PALUTIKOF, J. P., VAN DER LINDEN, P. J. & HANSON, C. E. 2007. *Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, 2007,* Cambridge, UK, Cambridge University Press.

PELAZ, S., DITTA, G. S., BAUMANN, E., WISMAN, E. & YANOFSKY, M. F. 2000. B and C floral organ identity functions require SEPALLATA MADS-box genes. *Nature,* 405**,** 200-3.

PERALES, M. & MAS, P. 2007. A functional link between rhythmic changes in chromatin structure and the Arabidopsis biological clock. *Plant Cell,* 19**,** 2111-2123.

PLESCH, G., EHRHARDT, T. & MUELLER-ROEBER, B. 2001. Involvement of TAAAG elements suggests a role for Dof transcription factors in guard cell-specific gene expression. *The Plant journal : for cell and molecular biology,* 28**,** 455-64.

POSE, D., YANT, L. & SCHMID, M. 2012. The end of innocence: flowering networks explode in complexity. *Curr Opin Plant Biol,* 15**,** 45-50.

PRICE, G. D., PENGELLY, J. J. L., FORSTER, B., DU, J. H., WHITNEY, S. M., VON CAEMMERER, S., BADGER, M. R., HOWITT, S. M. & EVANS, J. R. 2013. The cyanobacterial CCM as a source of genes for improving photosynthetic CO2 fixation in crop species. *Journal of Experimental Botany,* 64**,** 753-768.

RAWAT, R., SCHWARTZ, J., JONES, M. A., SAIRANEN, I., CHENG, Y., ANDERSSON, C. R., ZHAO, Y., LJUNG, K. & HARMER, S. L. 2009. REVEILLE1, a Myb-like transcription factor, integrates the circadian clock and auxin pathways. *Proceedings of the National Academy of Sciences,* 106**,** 16883-16888.

REPPERT, S. M. 2000. Cellular and molecular basis of circadian timing in mammals. *Semin Perinatol,* 24**,** 243-6.

RIECHMANN, J. L., HEARD, J., MARTIN, G., REUBER, L., JIANG, C., KEDDIE, J., ADAM, L., PINEDA, O., RATCLIFFE, O. J., SAMAHA, R. R., CREELMAN, R., PILGRIM, M., BROUN, P., ZHANG, J. Z., GHANDEHARI, D., SHERMAN, B. K. & YU, G. 2000. Arabidopsis transcription factors: genome-wide comparative analysis among eukaryotes. *Science,* 290**,** 2105-10.

ROCCARO, M. & SOMSSICH, I. E. 2011. Chromatin immunoprecipitation to identify global targets of WRKY transcription factor family members involved in plant immunity. *Methods Mol Biol,* 712**,** 45-58.

ROSINSKI, J. A. & ATCHLEY, W. R. 1998. Molecular evolution of the Myb family of transcription factors: evidence for polyphyletic origin. *J Mol Evol,* 46**,** 74-83.
SAGAR, M., CHERVIN, C., MILA, I., HAO, Y., ROUSTAN, J. P., BENICHOU, M., GIBON, Y., BIAIS, B., MAURY, P., LATCHE, A., PECH, J. C., BOUZAYEN, M. & ZOUINE, M. 2013. SlARF4, an auxin response factor involved in the control of sugar metabolism during tomato fruit development. *Plant Physiol,* 161**,** 1362-74.
SAGE, R. F., SAGE, T. L. & KOCACINAR, F. 2012. Photorespiration and the evolution of C4 photosynthesis. *Annu Rev Plant Biol,* 63**,** 19-47.
SALOME, P. A. & MCCLUNG, C. R. 2005. PSEUDO-RESPONSE REGULATOR 7 and 9 are partially redundant genes essential for the temperature responsiveness of the arabidopsis circadian clock. *Plant Cell,* 17**,** 791-803.
SALOME, P. A., TO, J. P. C., KIEBER, J. J. & MCCLUNG, C. R. 2006. Arabidopsis response regulators ARR3 and ARR4 play cytokinin-independent roles in the control of circadian period. *Plant Cell,* 18**,** 55-69.
SANCHEZ, A., SHIN, J. & DAVIS, S. J. 2011. Abiotic stress and the plant circadian clock. *Plant Signal Behav,* 6**,** 223-31.
SANTOPOLO, S., BOCCACCINI, A., LORRAI, R., RUTA, V., CAPAUTO, D., MINUTELLO, E., SERINO, G., COSTANTINO, P. & VITTORIOSO, P. 2015. DOF AFFECTING GERMINATION 2 is a positive regulator of light-mediated seed germination and is repressed by DOF AFFECTING GERMINATION 1. *Bmc Plant Biology,* 15.
SASAKI, N., MATSUMARU, M., ODAIRA, S., NAKATA, A., NAKATA, K., NAKAYAMA, I., YAMAGUCHI, K. & NYUNOYA, H. 2015. Transient expression of tobacco BBF1-related Dof proteins, BBF2 and BBF3, upregulates genes involved in virus resistance and pathogen defense. *Physiological and Molecular Plant Pathology,* 89**,** 70-77.
SCHMIED, K. C. & MERKLE, T. 2005. *A small family of LHY-CCA1-like (LCL) MYB1R transcription factors: potential co-regulators of the circadian oscillator.* [Online].
SCHUTZE, K., HARTER, K. & CHABAN, C. 2008. Post-translational regulation of plant bZIP factors. *Trends in plant science,* 13**,** 247-55.
SHIGYO, M., TABEI, N., YONEYAMA, T. & YANAGISAWA, S. 2007. Evolutionary processes during the formation of the plant-specific Dof transcription factor family. *Plant Cell Physiol,* 48**,** 179-85.
SILVERA, K., SANTIAGO, L. S., CUSHMAN, J. C. & WINTER, K. 2009. Crassulacean Acid Metabolism and Epiphytism Linked to Adaptive Radiations in the Orchidaceae. *Plant Physiology,* 149**,** 1838-1847.
SIPES, D. L. & TING, I. P. 1985. Crassulacean Acid Metabolism and Crassulacean Acid Metabolism Modifications in Peperomia camptotricha. *Plant Physiol,* 77**,** 59-63.
SOMERS, D. E., DEVLIN, P. F. & KAY, S. A. 1998. Phytochromes and cryptochromes in the entrainment of the Arabidopsis circadian clock. *Science,* 282**,** 1488-90.
SOMERVILLE, C., YOUNGS, H., TAYLOR, C., DAVIS, S. C. & LONG, S. P. 2010. Feedstocks for Lignocellulosic Biofuels. *Science,* 329**,** 790-792.
SONG, Y. H., ESTRADA, D. A., JOHNSON, R. S., KIM, S. K., LEE, S. Y., MACCOSS, M. J. & IMAIZUMI, T. 2014. Distinct roles of FKF1, GIGANTEA, and ZEITLUPE proteins in the regulation of CONSTANS stability in Arabidopsis photoperiodic flowering. *Proceedings of the National Academy of Sciences of the United States of America,* 111**,** 17672-17677.

SUZUKI, M., YAMAGUCHI, N., WATANABE, H., MORISHITA, T., MATSUYAMA, Y. & KOMEDA, Y. 2006. Genetic regulation of the floral integrators in Arabidopsis thaliana. *Plant and Cell Physiology,* 47**,** S2-S2.

TANAKA, M., TAKAHATA, Y., NAKAYAMA, H., NAKATANI, M. & TAHARA, M. 2009. Altered carbohydrate metabolism in the storage roots of sweet potato plants overexpressing the SRF1 gene, which encodes a Dof zinc finger transcription factor. *Planta,* 230**,** 737-46.

TAYBI, T., NIMMO, H. G. & BORLAND, A. M. 2004. Expression of phosphoenolpyruvate carboxylase and phosphoenolpyruvate carboxylase kinase genes. Implications for genotypic cavacity and phenotypic plasticity in the expression of crassulacean acid metabolism. *Plant Physiology,* 135**,** 587-598.

TAYBI, T., PATIL, S., CHOLLET, R. & CUSHMAN, J. C. 2000. A minimal serine/threonine protein kinase circadianly regulates phosphoenolpyruvate carboxylase activity in crassulacean acid metabolism-induced leaves of the common ice plant. *Plant Physiology,* 123**,** 1471-1481.

TEPER-BAMNOLKER, P. & SAMACH, A. 2005. The flowering integrator FT regulates SEPALLATA3 and FRUITFULL accumulation in Arabidopsis leaves. *Plant Cell,* 17**,** 2661-75.

TEPPERMAN, J. M., ZHU, T., CHANG, H. S., WANG, X. & QUAIL, P. H. 2001. Multiple transcription-factor genes are early targets of phytochrome A signaling. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 9437-42.

TIAN, Q., UHLIR, N. J. & REED, J. W. 2002. Arabidopsis SHY2/IAA3 inhibits auxin-regulated gene expression. *Plant Cell,* 14**,** 301-19.

TSUKAGOSHI, H., SUZUKI, T., NISHIKAWA, K., AGARIE, S., ISHIGURO, S. & HIGASHIYAMA, T. 2015. RNA-seq analysis of the response of the halophyte, Mesembryanthemum crystallinum (ice plant) to high salinity. *PloS one,* 10**,** e0118339.

UMEMURA, Y., ISHIDUKA, T., YAMAMOTO, R. & ESAKA, M. 2004. The Dof domain, a zinc finger DNA-binding domain conserved only in higher plants, truly functions as a Cys2/Cys2 Zn finger domain. *Plant J,* 37**,** 741-9.

VENKATESH, J. & PARK, S. W. 2015. Genome-wide analysis and expression profiling of DNA-binding with one zinc finger (Dof) transcription factor family in potato. *Plant Physiol Biochem,* 94**,** 73-85.

VICENTE-CARBAJOSA, J., MOOSE, S. P., PARSONS, R. L. & SCHMIDT, R. J. 1997. A maize zinc-finger protein binds the prolamin box in zein gene promoters and interacts with the basic leucine zipper transcriptional activator Opaque2. *Proceedings of the National Academy of Sciences of the United States of America,* 94**,** 7685-90.

VINCENTZ, M., BANDEIRA-KOBARG, C., GAUER, L., SCHLOGL, P. & LEITE, A. 2003. Evolutionary pattern of angiosperm bZIP factors homologous to the maize Opaque2 regulatory protein. *Journal of Molecular Evolution,* 56**,** 105-116.

VISION, T. J., BROWN, D. G. & TANKSLEY, S. D. 2000b. The origins of genomic duplications in Arabidopsis. *Science,* 290**,** 2114-7.

VISION, T. J., BROWN, D. G., SHMOYS, D. B., DURRETT, R. T. & TANKSLEY, S. D. 2000a. Selective mapping: a strategy for optimizing the construction of high-density linkage maps. *Genetics,* 155**,** 407-20.

WANG, Z. Y. & TOBIN, E. M. 1998. Constitutive expression of the CIRCADIAN CLOCK ASSOCIATED 1 (CCA1) gene disrupts circadian rhythms and suppresses its own expression. *Cell,* 93**,** 1207-1217.

WARREN, D. M. & WILKINS, M. B. 1961. Endogenous Rhythm in Rate of Dark-Fixation of Carbon Dioxide in Leaves of Bryophyllum-Fedtschenkoi. *Nature,* 191**,** 686-691.
WASHIO, K. 2001. Identification of Dof proteins with implication in the gibberellin-regulated expression of a peptidase gene following the germination of rice grains. *Bba-Gene Struct Expr,* 1520**,** 54-62.
WELLMER, F. & RIECHMANN, J. L. 2010. Gene networks controlling the initiation of flower development. *Trends in Genetics,* 26**,** 519-527.
WELTMEIER, F., EHLERT, A., MAYER, C. S., DIETRICH, K., WANG, X., SCHUTZE, K., ALONSO, R., HARTER, K., VICENTE-CARBAJOSA, J. & DROGE-LASER, W. 2006. Combinatorial control of Arabidopsis proline dehydrogenase transcription by specific heterodimerisation of bZIP transcription factors. *EMBO J,* 25**,** 3133-43.
WILKINS, M. B. 1959. AN ENDOGENOUS RHYTHM IN THE RATE OF CARBON DIOXIDE OUTPUT OF BRYOPHYLLUM .1. SOME PRELIMINARY EXPERIMENTS. *Journal of Experimental Botany,* 10**,** 377-390.
WILKINS, M. B. 1962. An endogenous rhythm in the rate of carbon dioxide output of Bryophyllum III. The effects of temperature on the phase and period of the rhythm. *Proc. R. Soc. London,* 156.
WILKINS, M. B. 1992. CIRCADIAN-RHYTHMS - THEIR ORIGIN AND CONTROL. *New Phytologist,* 121**,** 347-375.
WINTER, K. & HOLTUM, J. A. 2007. Environment or development? Lifetime net CO2 exchange and control of the expression of Crassulacean acid metabolism in Mesembryanthemum crystallinum. *Plant physiology,* 143**,** 98-107.
WINTER, K. & SMITH, J. A. C. 1996a. Crassulacean acid metabolism: Current status and perspectives. *Crassulacean Acid Metabolism,* 114**,** 389-426.
WINTER, K. & SMITH, J. A. C. 1996b. An introduction to crassulacean acid metabolism. Biochemical principles and ecological diversity. *In:* WINTER, K. & SMITH, J. A. C. (eds.) *Crassulacean Acid Metabolism - Biochemistry, Ecophysiology and Evolution.* Berlin 33: Springer-Verlag Berlin.
WU, C. Y., YOU, C. J., LI, C. S., LONG, T., CHEN, G. X., BYRNE, M. E. & ZHANG, Q. F. 2008. RID1, encoding a Cys2/His2-type zinc finger transcription factor, acts as a master switch from vegetative to floral development in rice. *Proceedings of the National Academy of Sciences of the United States of America,* 105**,** 12915-12920.
WU, X., DINNENY, J. R., CRAWFORD, K. M., RHEE, Y., CITOVSKY, V., ZAMBRYSKI, P. C. & WEIGEL, D. 2003. Modes of intercellular transcription factor movement in the Arabidopsis apex. *Development,* 130**,** 3735-45.
YANAGISAWA, S. 1995. A novel DNA-binding domain that may form a single zinc finger motif. *Nucleic acids research,* 23**,** 3403-10.
YANAGISAWA, S. 1997. Dof DNA-binding domains of plant transcription factors contribute to multiple protein-protein interactions. *European journal of biochemistry / FEBS,* 250**,** 403-10.
YANAGISAWA, S. 2000. Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. *Plant Journal,* 21**,** 281-288.
YANAGISAWA, S. 2002. The Dof family of plant transcription factors. *Trends in Plant Science,* 7**,** 555-60.

YANAGISAWA, S. 2004. Dof domain proteins: plant-specific transcription factors associated with diverse phenomena unique to plants. *Plant Cell Physiol,* 45**,** 386-91.

YANAGISAWA, S., AKIYAMA, A., KISAKA, H., UCHIMIYA, H. & MIWA, T. 2004. Metabolic engineering with Dof1 transcription factor in plants: Improved nitrogen assimilation and growth under low-nitrogen conditions. *Proceedings of the National Academy of Sciences of the United States of America,* 101**,** 7833-8.

YANAGISAWA, S. & IZUI, K. 1993. Molecular cloning of two DNA-binding proteins of maize that are structurally different but interact with the same sequence motif. *The Journal of biological chemistry,* 268**,** 16028-36.

YANAGISAWA, S. & SHEEN, J. 1998. Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. *Plant Cell,* 10**,** 75-89.

YANG, J., YANG, M. F., WANG, D., CHEN, F. & SHEN, S. H. 2010. JcDof1, a Dof transcription factor gene, is associated with the light-mediated circadian clock in Jatropha curcas. *Physiol Plant,* 139**,** 324-34.

YANG, X. H., CUSHMAN, J. C., BORLAND, A. M., EDWARDS, E. J., WULLSCHLEGER, S. D., TUSKAN, G. A., OWEN, N. A., GRIFFITHS, H., SMITH, J. A. C., DE PAOLI, H. C., WESTON, D. J., COTTINGHAM, R., HARTWELL, J., DAVIS, S. C., SILVERA, K., MING, R., SCHLAUCH, K., ABRAHAM, P., STEWART, J. R., GUO, H. B., ALBION, R., HA, J. M., LIM, S. D., WONE, B. W. M., YIM, W. C., GARCIA, T., MAYER, J. A., PETEREIT, J., NAIR, S. S., CASEY, E., HETTICH, R. L., CEUSTERS, J., RANJAN, P., PALLA, K. J., YIN, H. F., REYES-GARCIA, C., ANDRADE, J. L., FRESCHI, L., BELTRAN, J. D., DEVER, L. V., BOXALL, S. F., WALLER, J., DAVIES, J., BUPPHADA, P., KADU, N., WINTER, K., SAGE, R. F., AGUILAR, C. N., SCHMUTZ, J., JENKINS, J. & HOLTUM, J. A. M. 2015. A roadmap for research on crassulacean acid metabolism (CAM) to enhance sustainable food and bioenergy production in a hotter, drier world. *New Phytologist,* 207**,** 491-504.

YOO, S. K., CHUNG, K. S., KIM, J., LEE, J. H., HONG, S. M., YOO, S. J., YOO, S. Y., LEE, J. S. & AHN, J. H. 2005. CONSTANS activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to promote flowering in Arabidopsis. *Plant physiology,* 139**,** 770-8.

ZARZYCKI, J., AXEN, S. D., KINNEY, J. N. & KERFELD, C. A. 2013. Cyanobacterial-based approaches to improving photosynthesis in plants. *Journal of Experimental Botany,* 64**,** 787-798.

ZEEVAART, J. A. D. 1976. Physiology of flower formation. *Annu. Rev. Plant Physiol.,* 27**,** 321-348.

ZHANG, J. Z. 2003. Overexpression analysis of plant transcription factors. *Curr Opin Plant Biol,* 6**,** 430-40.

ZHANG, B., CHEN, W., FOLEY, R. C., BUTTNER, M. & SINGH, K. B. 1995. Interactions between distinct types of DNA binding proteins enhance binding to ocs element promoter sequences. *Plant Cell,* 7**,** 2241-2252.

ZHANG, Z. B., ZHU, J., GAO, J. F., WANG, C., LI, H., LI, H., ZHANG, H. Q., ZHANG, S., WANG, D. M., WANG, Q. X., HUANG, H., XIA, H. J. & YANG, Z. N. 2007. Transcription factor AtMYB103 is required for anther development by regulating tapetum development, callose dissolution and exine formation in Arabidopsis. *The Plant journal : for cell and molecular biology,* 52**,** 528-38.

ZOU, X. L., NEUMAN, D. & SHEN, Q. J. 2008. Interactions of two transcriptional repressors and two transcriptional activators in modulating gibberellin signaling in aleurone cells. *Plant Physiology,* 148**,** 176-186.

efficient dark CO_2 fixation, lower starch and malate levels, a reduction in the K_i of PEPC for malate, and also led to lower productivity under drought-stress. For CAM to function efficiently, *KfMYB439* needs to be strictly regulated, and the findings presented here suggest that this is achieved through this gene functioning through close interactions with the core oscillator.

The results presented here also support the original proposal that *KfCIB1* also functions as part of a signal transduction pathway between the circadian clock and CAM. An ortholog of *KfCIB1* was originally discovered in *M. crystallinum* (*McCIB1*) by virtue of the fact that its transcript level was up-regulated coincident with CAM induction, and, once induced, it oscillated with a robust diurnal and circadian rhythm. The orthologous gene in *K. fedtschenkoi* was also discovered to be a CAM-induced and clock-controlled TF, and previous work by PhD student Roland Meszter in the Hartwell lab had demonstrated that a heterologously expressed version of *KfCIB1* could bind to putative ACGT-motifs within the *KfPPCK1* promoter (Meszter, 2010). In this thesis, when *KfCIB1* was mis-expressed and plants were assayed under LL conditions, the regulation of the key clock genes *KfCCA1* and *KfTOC1* showed significant changes in the timing of their transcript oscillations. Taken together the LD and LL transcript oscillation changes suggest that *KfCIB1* not only plays a role in the output pathway of the circadian clock, but also is able to feed back to the core oscillator. *KfCIB1* also had a major effect on stomatal aperture and its regulation, with transgenic lines showing large and rapid changes in stomatal conductance as light levels changed between light and dark, which resulted in large but very short bursts of the release and fixation of CO_2 at these times. Mis-regulation of *KfCIB1* had some positive effects on *K. fedtschenkoi*, as the RNAi lines showed modest increases in

dark CO_2 fixation, the over-expresser lines also had an increase in malate levels, and starch levels accumulated to higher concentrations during the light period. Furthermore, in drought conditions both the over-expresser and RNAi knockdown lines showed an increase in dry weight matter compared to wild type, indicating that they were able to achieve more productive growth when water was limited.

KfCDF2 was the third CAM-induced and clock-controlled TF studied in this PhD. This TF was found to not only be involved in the circadian control of CAM, but also in the circadian control of other processes like flowering. In LL conditions, perturbed expression of *KfCDF2* was shown to impact on the expression level of *KfCCA1, KfTOC1* and *KfPPCK1*, with *FL_13A, FL_14C and RNAi_26B* all showing decreases at peak times of expression. *KfCDF2* was shown further to be involved in the circadian clock output pathways, as CO_2 exchange in LL was found to collapse to arrhythmia after a few days, whilst wild type continued to oscillate robustly. The transgenic changes to *KfCDF2* expression levels had mostly detrimental effects on CAM. Less malate was accumulated, likely due to a reduction in the K_i of PEPC for malate. Interestingly, under drought-stress, *KfCDF2* mis-regulation was shown to lead to an increase in succulence. After 31 days of no water, all *KfCDF2* transgenic lines, both overexpresser and RNAi knockdown, possessed a higher percentage moisture content throughout the whole plant compared to the wild type. When plants were drought-stressed for 90 days, all *KfCDF2* transgenic lines showed a higher level of succulence compared to the wild type. Also, in well-watered conditions the *KfCDF2* lines were able to increase succulence more rapidly than wild type, which reached a certain succulence level, and then maintained it. Furthermore, the *KfCDF2* over-expresser lines were shown to be involved in the circadian control of other processes

beyond just CAM, with the *KfCDF2_FL* lines displaying photoperiod insensitive flowering. Under certain day-length conditions, the *KfCDF2_RNAi* lines also attempted to initiate flowering at a faster rate than wild type. From screening the transcript abundance of flowering time genes in *K. fedtschenkoi*, there were many expression changes in the lines that had *KfCDF2* mis-expression, which could potential indicate the signalling pathway that was causing early flowering in these lines. Changes in *KfCDF2* levels also led to the plants becoming more sensitive to the quantity or quality of light, which suggests that this TF could play a role in the light input pathway to the clock.

Almost all of the TF transgenic *K. fedtschenkoi* lines studied in this thesis were able to produce more adventitious leaf margin plantlets than the wild type, which may be a consequence of CAM not functioning efficiently in the mature leaves that form the plantlets along their margin, which would potentially allow a greater availability of resources (carbohydrates, ATP, reducing equivalents etc) meaning that the transgenic lines could put more effort into reproduction, which may be an advantage if the leaves are not able to sustain themselves using the CAM pathway (i.e. rather than the mother leaf surviving, the transgenic lines divert more of their resources into reproduction in the form of adventitious plantlets).

8.2 Future directions for research into the genes responsible for the circadian control of CAM

The results presented in this thesis not only suggest many more interesting avenues of research for *KfMYB439, KfCIB1* and *KfCDF2*, especially the identification of their target motifs and the downstream genes they regulate, but also demonstrate that the original approach of using transcriptome sequencing data to identify potential genes

involved in CAM and its circadian regulation was effective. This supports the further investigation of other regulatory proteins identified from the currently available *K. fedtschenkoi* SOLiD and Illumina RNA-seq datasets. In part of this PhD, some time was spent on the identification of new genes that could be involved in CAMs circadian optimisation. One gene identified was an ortholog of the Arabidopsis gene *SHORT HYPOCOTYL 2* (*SHY2*), which was found to peak at the end of the dark period, like *KfPPCK1*, and also had a CAM:C_3 ratio of 251, showing strong up-regulation in CAM tissues (Fig. 8.1A). Using a further SOLiD RNA-seq dataset that analysed a range of different tissues and organs, including roots, aerial roots, stems, leaf epidermal peels and flower buds, transcripts of this gene were also discovered to be most abundant in epidermal peels. This could therefore suggest that this gene may be a CAM-associated transcription factor that functions in the unique inverse pattern of stomatal opening and closing characteristic of CAM species. *SHY2* in Arabidopsis regulates auxin responses in roots (Tian *et al.,* 2002), but was not detected in root samples in the *K. fedtschenkoi* RNA-seq data, except for during the dark period in aerial roots (Fig. 8.1B). This therefore supports the proposal that *KfSHY2* may have been recruited to a new function in the regulation of stomata during CAM in *K. fedtschenkoi*.

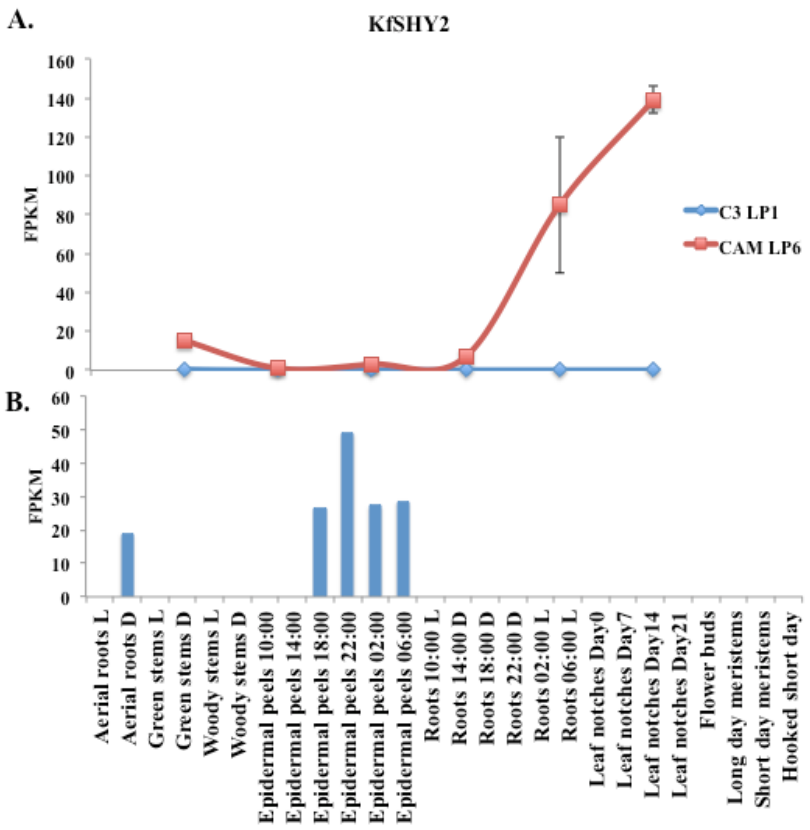

Figure 8.1. Quantitative RNA-seq analysis enables the identification of a novel, diurnally oscillating transcription factor in *K. fedtschenkoi* that is orthologous to *SHY2*. **A.** Diurnal control of transcript abundance (FPKM) of *KfSHY2*, a novel CAM-induced gene. **B**. SOLiD RNA-seq data showing transcript abundance of *KfSHY2* in different locations throughout the plant, showing that this gene was most highly expressed in epidermal peels, suggesting possible roles in stomatal control.

Another gene identified was *Contig6548* (Fig. 8.2). This gene was found to have no known homolog in Arabidopsis, but had a CAM:C_3 transcript abundance ratio of 132 from the SOLiD RNA-seq data for C_3 and CAM leaves, and reached its highest transcript abundance at the end of the dark period (Fig. 8.2A). This could therefore suggest that a novel gene specific to *K. fedtschenkoi,* which is not present in Arabidopsis, has evolved to function in CAM regulation. Transcripts for this gene were also found to most abundant in epidermal peels, especially in the dark, which would again suggest that this gene may be involved in stomatal control during CAM (Fig. 8.2B).

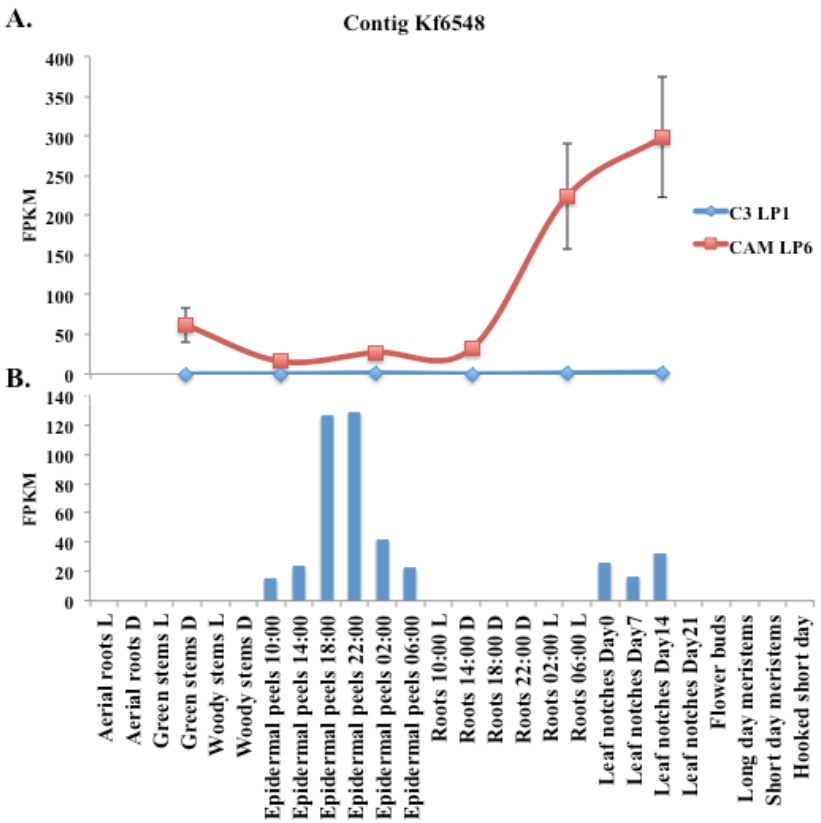

Figure 8.2. Quantitative RNA-seq analysis enables the identification of a novel, diurnally oscillating transcription factor in *K. fedtschenkoi* with no known ortholog in Arabidopsis **A.** Diurnal control of transcript abundance (FPKM) of *Contig6548*, a novel CAM-induced gene. **B.** SOLiD RNA-seq data showing transcript abundance of *Contig6548* in different locations throughout the plant, showing that this gene was most highly expressed in epidermal peels, suggesting possible roles in stomatal control.

Many more genes can be identified in this way through in depth analysis of the existing RNA-seq datasets for *K. fedtschenkoi*. The manipulation of these other novel putative regulators of CAM in transgenic lines of *K. fedtschenkoi* or *K. laxiflora* may reveal their role, if any, in the developmental induction and daily regulation of CAM. However, the results presented in this thesis suggest that there is a good chance that many of these novel genes will not show large changes in malate or starch level, as seen with *KfMYB439, KfCIB1* and *KfCDF2*. Initial screening of these transgenic lines may in the future prove more fruitful if quantitative malate and starch assays are

performed rather than doing iodine and chlorophenol red screens, which for the lines discussed in this thesis, did not show consistent or clear results.

It should be noted though, that throughout the interrogation of circadian regulation of genes, both in LD and especially in LL conditions, there were problems with the consistency of WT rhythms. This is likely due to the time of year different batches of plants were grown within the rooftop greenhouses, resulting in changes to amount of natural light they received. This could not be avoided though, as this was the only area with a big enough amount of space available to grow these lines. Simultaneous growth would have undoubtedly been the best solution, but there was not enough space available to conduct all of the sampling experiments at the same time in controlled conditions. Further work in this area should be undertaken to try and understand why the rhythms are slightly different depending on the time of year grown, even after being entrained for 1 week in climate controlled growth chambers with the exact same conditions. It would be interesting to see if longer entrainment before sampling could alter this issue.

Furthermore, issues with the rhythmicity of some genes investigated in LL conditions are apparent. Throughout all the LL experiments only the gene *KfPPCK* oscillates robustly, whilst for *KfMYB439. KfCIB1, KfCDF2, KfTOC1* and *KfCCA1* rhythms are hard to define in most timecourses. Software has been designed for this purpose, such as Biological Rhythm Analysis Software Suite (BRASS), and its improved software Biological Data Repository (BioDare), which are able to provide information about rhythmic properties within an experimental time-series (Zielinski *et al.,* 2014). Although this would be a highly useful tool, and already is to many circadian groups globally, the data collected in this PhD thesis did not meet the requirements to enable

its use. A large number of repeats are required for this software, whilst due to time, space and resource constraints, only 3 were able to be collected for both the LD and LL timecourses. Furthermore, the longer the rhythm is measured for the more accurate results, with 5 days or longer being suggested as the ideal length of time, with samples taken every 2hours (Zielinski *et al.,* 2014). Again this would be unfeasible with the equipment and space available at the University of Liverpool, especially as *K. fedtschenkoi* take over 6 months to grow to a suitable size for sampling. This is why techniques such as luciferase imaging or delayed fluorescence are readily used to assess rhythmicity, as large amounts of circadian data can be collected this way, with little effort, as cameras can be used in growth chambers (Tindall *et al.,* 2015).

8.3 Future Work for *KfMYB439*, *KfCIB1* and *KfCDF2*

One key area for future research on all three of the TFs investigated in this PhD will be to identify their target sites within the *K. fedtschenkoi* or *K. laxiflora* genome, to aid in understanding the genes that have their expression regulated by *KfMYB439, KfCIB1* and *KfCDF2*. This would also enable identification of genes likely regulated in an indirect way too. A powerful approach for discovering the DNA-motifs bound by these TFs is Chromatin Immunoprecipitation followed by sequencing (ChIP-Seq). Over-expression of each TF fused with an immuno-affinity tag should facilitate the preferential capture TF-DNA complexes using a specific antibody for the tag attached to the TF. The TFs are removed and the remaining DNA can be sequenced using the Illumina system and the resulting reads are mapped to the annotated genomic scaffolds. Transgenic lines of *K. laxiflora* have already been created for this ChIP-seq work with *KfMYB439, KfCIB1* and *KfCDF2*. A Haemagglutenin- (HA-) tag was used,

which has previously been shown to be effective in ChIP-Seq for plants (Kaufmann *et al.,* 2010; Roccaro *et al.,* 2011; Miao *et al.,* 2013). Two different HA-tag constructs were generated for each TF. One had an HA-tag attached to the N-terminus of the full length ORF of the TF, and the other had the HA-tag attached to the C-terminus. The HA-tag was fused to both ends of the protein to control for any problems with protein folding affecting the accessibility of the tag for the antibody. These transgenic lines over-express the *K. fedtschenkoi* gene, but in the closely related species *K. laxiflora*. *K. laxiflora* was chosen for these experiments as it is able to set viable seed, whereas *K. fedtschenkoi* fails to do this. The genome of *K. laxiflora* is also available in a draft version, allowing for the identification of the gene promoters bound by each of the TFs (Yang *et al.,* 2015).

The HA-tag transgenic lines were created as described in materials and methods (Chapter 2.1.5.2). Screening was then conducted, which again found that the over-expression of *KfCDF2* led to photoperiod independent flowering and increased succulence. This demonstrated that the HA-tag did not interfere with the normal function of this protein as the phenotypes matched those reported for the over-expresser transgenic lines of *K. fedtschenkoi* in which *KfCDF2* was up-regulated constitutively. To date, only transgene transcript levels have been tested for these new HA-tag transgenic lines of *K. laxiflora* (Fig. 8.3). One set of internal primers for the gene was used, which amplified both the *K. laxiflora* endogenous transcript and the HA-tagged transgene (Fig. 8.3A, C, and E). In addition, a second set of primers was used for which one primer targeted the HA-tag (Fig. 8.3B,D, and F), and so would only amplify mRNA expressed from the transgene that included the HA-tag.

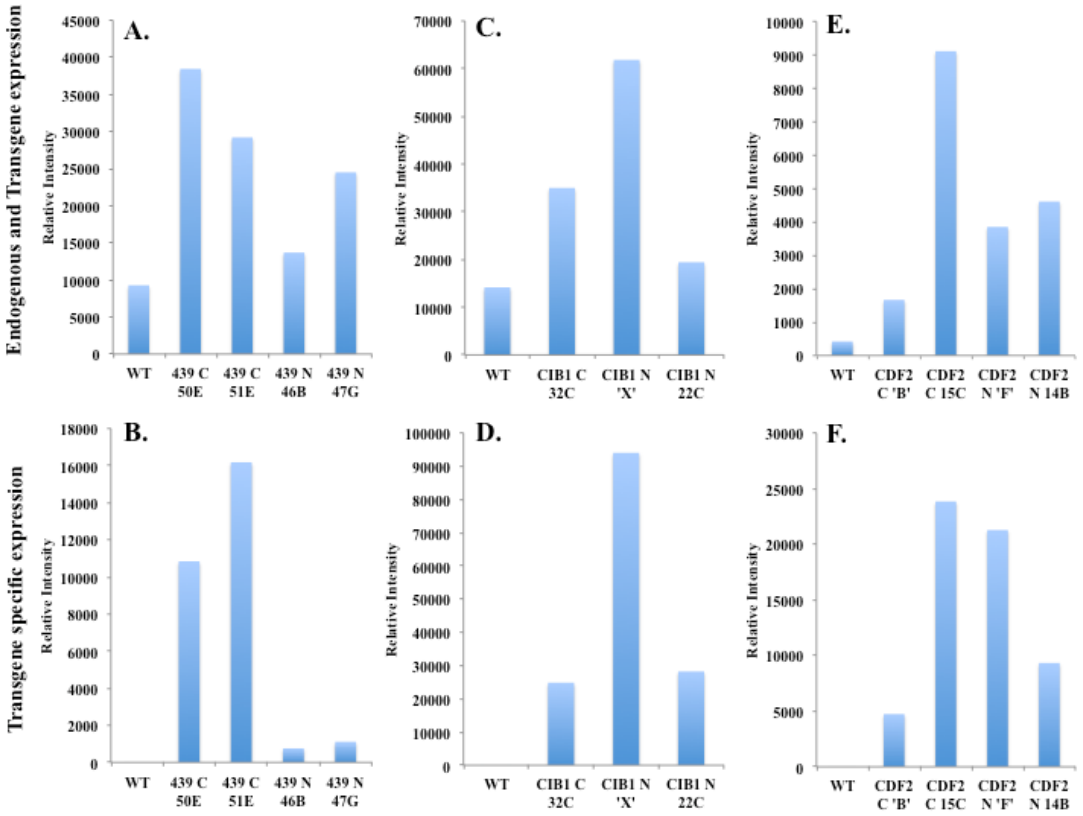

Figure 8.3. Transgene transcript abundance for the HA-tagged over-expresser lines of *K. laxiflora* containing HA- tagged version of the full length open reading frame for *KfMYB439, KfCIB1* and *KfCDF2*. Transcript levels were determined using LP6 samples collected at 22:00 (2h before dawn) for both the endogenous and transgene: **A.** *KfMYB439*, **C.** *KfCIB1*, **E.** *KfCDF2*. Only HA-tagged transgene expression: **B.** *HA-KfMYB439*, **D.** *HA-KfCIB1*, **F.** *HA-KfCDF2*.

From these initial screening results, lines were chosen for time courses to be taken for the ChIP-Seq experiment, based on the fact that they expressed more overall of the desired gene than the wild type, and also expressed the HA-tagged version (Fig. 8.3). These time courses were all collected and stored in the -80˚C freezer, but time ran out before the ChIP-seq experiment could be attempted.

Another interesting line of future work would be to create GFP-tagged transgenic lines for each of the TFs. This would enable protein localisation to be determined (Chalfie *et al.,* 1994). This would enable firstly detection of the protein, revealing how long after peak transcript expression the protein accumulates, and would also

reveal whether these genes are localised to specific cell types. GFP-tagged lines would also enable the characterisation of the temporal pattern of sub-cellular localisation for these TFs (e.g. when do they move to the nucleus so that they can bind to their target DNA), and the dynamics of their degradation (Chalfie *et al.,* 1994; Nakajima *et al.,* 2001; Wu *et al.,* 2003). These more detailed levels of information would be of particular interest for the further investigation of the function of *KfCDF2* in the regulation of flowering in response to short days, as the GFP-tagging would allow for the visualisation of where and when KfCDF2 was functioning within the different organs, tissues and cell types involved in the photoperiodic induction of flowering.

Finally, bZIP TFs can only bind to DNA as a dimer, single-MYB-like transcription factors have not be found to function as monomers, and DOFs interact with a variety of other proteins including bZIPs and MYB-like TFs. We have seen over-arching phenotypes between the three *K. fedtschenkoi* genes, such as many showing reductions in dark CO2 fixation, collapsed to arrhythmia for CO2 fixation in LL, and all showed a reduced Ki of PEPC for malate. This could therefore suggest interaction, or that they function within the same pathway. It is therefore important to identify the protein interaction partners of each of the three TFs studied here.

The powerful way to approach this question would be to identify other genes expressed at the same time as each TF, which encode proteins that show sequence homology, especially in areas enabling protein-protein interactions (Ahn *et al.,* 2007). A tagged version of the original TF can then be created, and incubated with the potential protein partner. Crosslinking and pull down of complexes can then be carried out, and electrophoretic mobility shift assays (EMSAs) can be conducted to

determine binding (Krohn *et al.,* 2002). Complexes can then be cut out of the gel and ran on an SDS page to separate out the proteins to confirm the heterodimer formation. Many protein-protein interactions have already been demonstrated using these techniques. These approaches would also be useful for *KfCIB1,* to see if it was able to form homodimers. Alternatively, the yeast two-hybrid system could also be utilised (Fields *et al.,* 1989).

To identify an unknown protein bound to one of the TFs is much more difficult, and methods for this have only recently been developed, using a combination of mass spectrometry and computer programmes that are able to determine the sequences of peptides. These approaches at present generate a lot of false positives (Ahn *et al.,* 2007). Improvements in this area are occurring constantly, and so in a few years it is likely that the identification of the protein interaction partners of KfMYB439, KfCIB1 and KfCDF2 will be much more readily achievable.

8.4 Final Thoughts

CAM is an adaptation of plants to arid environments, which already make up 40% of land on Earth (Borland *et al.,* 2009). Modelling predictions from climate change scientists suggest that this figure is only going to increase. The global population is also increasing, with humans predicted to rise from the 7 billion on Earth today to over 9 billion by 2050. These factors further compound the already large problems of global food and energy security.

CAM plants have great potential to act as a key part of humanity's response to the challenges we will face. They are not only able to survive in seasonally dry environments, but are in fact able to grow at near maximum productivity in water-limited environments (Nobel 1996; Borland *et al.,* 2009; Somerville *et al.,* 2010;

Borland *et al.,* 2011). Knowing the natural capabilities of plants that have pre-evolved improvements to the pathway of photosynthetic CO_2 fixation has led many research groups throughout the world to investigate ways to enhance photosynthesis as a means to improve agricultural production and sustainability (Covshoff *et al.,* 2012; Ducat *et al.,* 2012; Maurino *et al.,* 2013) Borland et al., 2014; Borland et al, 2015). Other recent projects including the engineering of C_4 photosynthesis into C_3 rice to enhance productivity (Covshoff *et al.,* 2012; Leegood 2013), introducing carbon-concentrating mechanisms (CCMs) from cyanobacteria into plants (Price *et al.,* 2013; Zarzycki *et al.,* 2013), and CAM into C_3 plants to improve crop water use efficiency (Yang *et al.,* 2015), are also currently being undertaken. For CAM to be engineered into C_3, the genes involved in its optimal temporal regulation need to be identified and characterised in detail, in addition to the identification of the genes/ proteins involved in the biochemical steps of the pathway (Borland et al., 2014; Yang et al., 2015). This PhD has revealed that there are many genes with potential functions in the daily coordination and optimisation of CAM in response to the central circadian clock. *KfMYB439* is a gene that must be regulated tightly in CAM plants in order to enable optimal CAM to occur, whilst *KfCIB1* may be involved in optimal control of stomatal responses to the light/ dark transition. Over-expression of *KfCDF2* led to the plants becoming more succulent, possibly at the cost of increased growth, but if CAM were only being engineered to be activated at times of water stress, then this gene could be a valuable tool for the development of plants that are better armed to survive prolonged drought-stress.

Rapid progress with improving crop drought tolerance is critical for the prevention of a global food and energy crisis in the future, and it is hoped that the research

presented in this PhD will make at least a small contribution to the ability of crop and plant scientist to produce more water use efficiency crops.